GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 3, 2001, 20:46:35 ; Search time 22.79 Seconds
(without alignments)
18.621 Million cell updates/sec

Title: US-09-377-081-18
Perfect score: 47
Sequence: 1 SCHLPWA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0601:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	47	100.0	7	21	AAY84191	Amino acid sequenc
2	47	100.0	22	22	AAB59936	Human leptin fragm
3	47	100.0	22	22	AAB59937	Human leptin fragm
4	47	100.0	22	22	AAB59938	Human leptin fragm
5	47	100.0	34	19	AAW45585	Peptide fragment o
6	47	100.0	54	17	AAW00304	Human ob protein s
7	47	100.0	54	17	AAW00047	Mammalian Ob prote
8	47	100.0	54	18	AAW17704	Human obese gene p
9	47	100.0	67	18	AAW17702	Human obese gene p
10	47	100.0	77	18	AAW34400	Human Met-OB prote
11	47	100.0	77	18	AAW27173	Human recombinant
12	47	100.0	105	17	AAR98746	Biologically activ
13	47	100.0	119	17	AAW02150	Anti-obesity prote
14	47	100.0	119	17	AAR92732	Obesity protein C-
15	47	100.0	119	21	AAB28458	Human OB protein C
16	47	100.0	119	21	AAB28477	Human OB protein C
17	47	100.0	119	21	AAY87736	Human OB protein f
18	47	100.0	130	19	AAW71300	Obesity protein an
19	47	100.0	138	17	AAW07434	Large monomer of m
20	47	100.0	144	17	AAW03524	Anti-obesity prote
21	47	100.0	145	17	AAW00302	Human delta Gln28
22	47	100.0	145	17	AAW00541	Human mature obesi
23	47	100.0	145	18	AAW30893	Synthetic obesity
24	47	100.0	145	21	AAB14266	Mature human lepti
25	47	100.0	145	21	AAY95787	Mature recombinant
26	47	100.0	145	21	AAY92815	Mature leptin rece
27	47	100.0	145	21	AAY97889	Mutant mature huma
28	47	100.0	145	21	AAY83769	Human OB mutein (V
29	47	100.0	146	17	AAW06524	Wild type ob prote
30	47	100.0	146	17	AAW00301	Human ob protein.
31	47	100.0	146	17	AAW00013	Acid stable modifi
32	47	100.0	146	17	AAR99490	Chimeric ob protei
33	47	100.0	146	17	AAR99493	Chimeric ob protei
34	47	100.0	146	17	AAR99489	Generic ob protein
35	47	100.0	146	17	AAR99498	Generic ob protein
36	47	100.0	146	17	AAR99499	Acid stable modifi
37	47	100.0	146	17	AAW00010	Acid stable modifi
38	47	100.0	146	17	AAW00012	Acid stable modifi
39	47	100.0	146	17	AAW00539	Human mature obesi
40	47	100.0	146	18	AAW30897	Synthetic obesity
41	47	100.0	146	18	AAW30892	Synthetic obesity
42	47	100.0	146	18	AAW30894	Synthetic obesity
43	47	100.0	146	18	AAW34483	Human obesity prot
44	47	100.0	146	18	AAW34489	Obesity protein an
45	47	100.0	146	21	AAY82111	Mature human obese

ALIGNMENTS

RESULT 1
AAY84191
ID AAY84191 standard; peptide; 7 AA.
XX
AC AAY84191;
XX
DT 03-JUL-2000 (first entry)
XX
DE Amino acid sequence of a peptide derived from human leptin.
XX
KW Human; leptin; blood brain barrier; homeostasis; body mass; anorexia;
KW obesity; hyperglycemia; hyperinsulinemia; hyperphagia;
KW thyroid dysfunction; infertility; type II diabetes mellitus;
KW non-insulin-dependent diabetes mellitus; hematopoiesis dysfunction;
KW tumour suppression; weight loss; diet.
XX
OS Homo sapiens.
XX
PN WO200011173-A1.
XX
PD 02-MAR-2000.
XX
PF 20-AUG-1999; 99WO-US19021.
XX
PR 21-AUG-1998; 98US-0097457.
PR 19-AUG-1999; 99US-0377081.
XX
PA (ALBA-) ALBANY MEDICAL COLLEGE.
XX
PI Grasso P, Lee DW, Leinung MC;
XX
DR WPI; 2000-237652/20.
XX
PT Leptin peptides useful for treating pathophysiology relating to

PT homeostasis of body mass such as obesity, anorexia, and hematopoiesis
PT dysfunction and tumor suppression -
XX
PS Claim 7; Page 79; 121pp; English.
XX
CC The present sequence represents a peptide derived from human leptin.
CC The specification describes leptin-derived peptides which have
CC increased ability to cross the blood brain barrier and improved
CC bio-availability. Peptides derived from leptin are useful for treating
CC and preventing pathophysiology relating to homeostasis of body mass
CC such as anorexia, obesity comprising hyperglycemia, hyperinsulinemia,
CC hyperphagia, thyroid dysfunction, infertility, type II diabetes mellitus
CC and non-insulin-dependent diabetes mellitus (NIDDM), and hematopoiesis
CC dysfunction and tumor suppression. The peptides are also useful for
CC identifying drugs useful in weight loss diet regimen.
XX
SQ Sequence 7 AA;

Query Match 100.0%; Score 47; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.4e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 1 schlpwa 7

RESULT 2
AAB59936
ID AAB59936 standard; Peptide; 22 AA.
XX
AC AAB59936;
XX
DT 06-JUN-2001 (first entry)
XX
DE Human leptin fragment SEQ ID NO: 54.
XX
KW Leptin; human; LSR; lipolysis stimulated receptor; obesity;
KW hypertension; anorexia; cachexia; stroke; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200121647-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-IB01470.
XX
PR 22-SEP-1999; 99US-0155506.
XX
PA (GEST) GENSET.
XX
PI Yen F, Erickson MR, Fruebis J, Bihain B;
XX
DR WPI; 2001-218642/22.
XX
PT New leptin polypeptide fragment and related polynucleotides, useful for
PT the prevention and treatment of obesity and obesity-related diseases
PT such as hypertension and diabetes -
XX
PS Example 10; Page 238; 247pp; English.
XX
CC The present invention provides the protein and coding sequences of leptin
CC fragments which modulate the activity of lipolysis stimulated factor
CC (LSR). These sequences are useful in the treatment of obesity related
CC diseases, including obesity, anorexia, cachexia, cardiac and coronary
CC insufficiency, stroke, hypertension, atheromatous disease,
CC atherosclerosis, non-insulin dependent diabetes, hyperlipidaemia,
CC hyperuricaemia and syndrome X.
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 47; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 15 schlpwa 21

RESULT 3
AAB59937
ID AAB59937 standard; Peptide; 22 AA.
XX
AC AAB59937;
XX
DT 06-JUN-2001 (first entry)
XX
DE Human leptin fragment SEQ ID NO: 55.
XX
KW Leptin; human; LSR; lipolysis stimulated receptor; obesity;
KW hypertension; anorexia; cachexia; stroke; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200121647-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-IB01470.
XX
PR 22-SEP-1999; 99US-0155506.
XX
PA (GEST) GENSET.
XX
PI Yen F, Erickson MR, Fruebis J, Bihain B;
XX
DR WPI; 2001-218642/22.
XX
PT New leptin polypeptide fragment and related polynucleotides, useful for
PT the prevention and treatment of obesity and obesity-related diseases
PT such as hypertension and diabetes -
XX
PS Example 10; Page 238; 247pp; English.
XX
CC The present invention provides the protein and coding sequences of leptin
CC fragments which modulate the activity of lipolysis stimulated factor
CC (LSR). These sequences are useful in the treatment of obesity related
CC diseases, including obesity, anorexia, cachexia, cardiac and coronary
CC insufficiency, stroke, hypertension, atheromatous disease,
CC atherosclerosis, non-insulin dependent diabetes, hyperlipidaemia,
CC hyperuricaemia and syndrome X.
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 47; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 10 schlpwa 16

RESULT 4
AAB59938
ID AAB59938 standard; Peptide; 22 AA.
XX
AC AAB59938;
XX
DT 06-JUN-2001 (first entry)
XX

DE Human leptin fragment SEQ ID NO: 56.
XX
KW Leptin; human; LSR; lipolysis stimulated receptor; obesity;
KW hypertension; anorexia; cachexia; stroke; atherosclerosis.
XX
OS Homo sapiens.
XX
PN WO200121647-A2.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-IB01470.
XX
PR 22-SEP-1999; 99US-0155506.
XX
PA (GEST) GENSET.
XX
PI Yen F, Erickson MR, Fruebis J, Bihain B;
XX
DR WPI; 2001-218642/22.
XX
PT New leptin polypeptide fragment and related polynucleotides, useful for
PT the prevention and treatment of obesity and obesity-related diseases
PT such as hypertension and diabetes -
XX
PS Example 10; Page 238; 247pp; English.
XX
CC The present invention provides the protein and coding sequences of leptin
CC fragments which modulate the activity of lipolysis stimulated factor
CC (LSR). These sequences are useful in the treatment of obesity related
CC diseases, including obesity, anorexia, cachexia, cardiac and coronary
CC insufficiency, stroke, hypertension, atheromatous disease,
CC atherosclerosis, non-insulin dependent diabetes, hyperlipidaemia,
CC hyperuricaemia and syndrome X.
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 47; DB 22; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 5 schlpwa 11

RESULT 5
AAW45585
ID AAW45585 standard; peptide; 34 AA.
XX
AC AAW45585;
XX
DT 03-JUN-1998 (first entry)
XX
DE Peptide fragment of leptin (ob 116-149) that modulates body weight.
XX
KW Leptin; obesity; body weight; diabetes; energy; metabolic disorder;
KW ob protein.
XX
OS Homo sapiens.
XX
PN WO9746585-A2.
XX
PD 11-DEC-1997.
XX
PF 04-JUN-1997; 97WO-EP02968.
XX
PR 20-FEB-1997; 97GB-0003493.
PR 06-JUN-1996; 96GB-0011775.
PR 05-SEP-1996; 96GB-0018540.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Albarazanji KA, Arch JR, Camilleri P, Neville WA;
XX
DR WPI; 1998-042120/04.
XX
PT Peptide fragments of leptin that modulate body weight by regulating
PT energy utilisation - especially useful for treatment of obesity and
PT diabetes
XX
PS Claim 4; Page 1; 19pp; English.
XX
CC The invention relates to specifically claimed peptides AAW45577-W45586
CC or their derivatives, analogues and variants, that modulate,
CC specifically reduce, body weight, mainly by affecting energy utilisation.
CC Also new are: (1) nucleic acid that encodes the peptides; (2) vectors
CC containing the nucleic acid; and (3) host cells transformed with this
CC vector. The peptides are used to treat nutritional or metabolic
CC disorders, particularly obesity and diabetes.
XX
SQ Sequence 34 AA;

Query Match 100.0%; Score 47; DB 19; Length 34;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 1 schlpwa 7

RESULT 6
AAW00304
ID AAW00304 standard; protein; 54 AA.
XX
AC AAW00304;
XX
DT 25-NOV-1996 (first entry)
XX
DE Human ob protein sequence.
XX
KW Human; anti-obesity; ob; adiposity regulating hormone; body fat;
KW obese; type II diabetes; cardiovascular disease; cancer; body weight;
KW cosmetics; animal feed additive.
XX
OS Homo sapiens.
XX
PN US5525705-A.
XX
PD 11-JUN-1996.
XX
PF 31-JAN-1995; 95US-0381370.
XX
PR 31-JAN-1995; 95US-0381370.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Dimarchi RD, Flora DB, Heath WF, Hoffmann JA, Shields JE;
PI Smiley DL;
XX
DR WPI; 1996-286451/29.
XX
PT New anti-obesity peptide(s) regulate fat tissue - are encoded by
PT human fat cell DNA and have improved stability
XX
PS Claim 8; Column -; 10pp; English.
XX
CC This sequence represents the human anti-obesity (ob) peptide.
CC Ob is thought to be an adiposity regulating hormone which regulates
CC body fat by a feedback model. When a mammal overeats the resulting
CC excess fat signals to the brain that the body is obese, which in turn
CC causes the body to eat less and burn more fuel. This protein is
CC biologically active for the treatment of obesity. Individuals treated

```
CC   with this protein have a reduced risk for type II diabetes,
CC   cardiovascular disease and cancer.  The ob protein can also be used
CC   to control body weight for cosmetic purposes and as an animal feed
CC   additive.
XX
SQ   Sequence   54 AA;


  Query Match             100.0%;  Score 47;  DB 17;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db       3 schlpwa 9


RESULT   7
AAW00047
ID   AAW00047 standard; peptide; 54 AA.
XX
AC   AAW00047;
XX
DT   23-OCT-1996  (first entry)
XX
DE   Mammalian Ob protein fragment #6.
XX
KW   Mammalian; ob protein; antibody; identification; quantitate;
KW   diagnosis; obese; obesity; peptide hormone.
XX
OS   Synthetic.
XX
PN   WO9623815-A1.
XX
PD   08-AUG-1996.
XX
PF   29-JAN-1996;   96WO-US00957.
XX
PR   31-JAN-1995;   95US-0381264.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Heath WF,  Manetta JV,  Shields JE;
XX
DR   WPI; 1996-371375/37.
XX
PT   Mono- and polyclonal antibodies specific for ob gene prods - useful
PT   to isolate and quantitate ob proteins in biological fluids derived
PT   from obese patients
XX
PS   Claim 1; Page 5; 45pp; English.
XX
CC   The sequences given in AAW00042-72 represent fragments of mammalian ob
CC   proteins which  may be used in the method of the invention to
CC   generate antibodies reactive with mammalian ob.  The antibodies are
CC   immobilised on a surface and used to isolate, identify and
CC   quantitate ob proteins from biological fluids.  The antibodies can
CC   be used to diagnose whether obese patients are expressing ob
CC   proteins.  Ob proteins are thought to be hormones which regulate the
CC   size of the body's fat depot by a feedback model.
XX
SQ   Sequence   54 AA;


  Query Match             100.0%;  Score 47;  DB 17;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db       3 schlpwa 9


RESULT   8
AAW17704
ID   AAW17704 standard; peptide; 54 AA.
XX
AC   AAW17704;
XX
DT   29-JAN-1998  (first entry)
XX
DE   Human obese gene product C-terminal fragment (amino acids 114-167).
XX
KW   Obese gene; ob; OBF; obese protein-associated diabetes; diagnosis;
KW   obesity; predisposition.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  4..54
FT                   /label= disulphide_bond
XX
PN   WO9716550-A1.
XX
PD   09-MAY-1997.
XX
PF   28-OCT-1996;   96WO-US17365.
XX
PR   02-NOV-1995;   95US-0007789.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Krystek SR,  Mapelli C,  Meyers CA,  Novotny J;
XX
DR   WPI; 1997-272119/24.
DR   P-PSDB; AAT68441.
XX
PT   Biologically active human and murine obese gene products - used to
PT   treat obese protein-associated diabetes or obesity
XX
PS   Claim 8; Page 2; 30pp; English.
XX
CC   The present sequence represents a novel C-terminal fragment of the human
CC   obese gene product (OBF). Novel biologically active C-terminal OBF
CC   fragments (AAW17701-4, murine and human) retain anti-obesity and/or
CC   anti-diabetic activities. The fragments contain less than 68 amino
CC   acids, two of which are capable of forming cross-linkages, preferably
CC   cysteines. OBF polypeptides can be used for treating diabetes, obesity
CC   or both in a patient in need of treatment, where the diabetes is obese
CC   protein-associated, and a concomitant decrease in food consumption does
CC   not occur. Labelled OBF can be used to diagnose obesity or a diabetes or
CC   a predisposition to develop obesity or diabetes.
XX
SQ   Sequence   54 AA;


  Query Match             100.0%;  Score 47;  DB 18;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 0.57;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db       3 schlpwa 9


RESULT   9
AAW17702
ID   AAW17702 standard; peptide; 67 AA.
XX
AC   AAW17702;
XX
DT   02-FEB-1998  (first entry)
XX
DE   Human obese gene product C-terminal fragment (amino acids 101-167).
XX
KW   Obese gene; ob; OBF; obese protein-associated diabetes; diagnosis;
```

```
KW   obesity; predisposition.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  17..67
FT                   /label= disulphide_bond
XX
PN   WO9716550-A1.
XX
PD   09-MAY-1997.
XX
PF   28-OCT-1996;    96WO-US17365.
XX
PR   02-NOV-1995;    95US-0007789.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Krystek SR,  Mapelli C,  Meyers CA,  Novotny J;
XX
DR   WPI; 1997-272119/24.
XX
PT   Biologically active human and murine obese gene products - used to
PT   treat obese protein-associated diabetes or obesity
XX
PS   Claim 6; Page 2; 30pp; English.
XX
CC   The present sequence represents a novel C-terminal fragment of the human
CC   obese gene product (OBF). Novel biologically active C-terminal OBF
CC   fragments (AAW17701-4, murine and human) retain anti-obesity and/or
CC   anti-diabetic activities. The fragments contain less than 68 amino
CC   acids, two of which are capable of forming cross-linkages, preferably
CC   cysteines. OBF polypeptides can be used for treating diabetes, obesity
CC   or both in a patient in need of treatment, where the diabetes is obese
CC   protein-associated, and a concomitant decrease in food consumption does
CC   not occur. Labelled OBF can be used to diagnose obesity or a diabetes or
CC   a predisposition to develop obesity or diabetes.
XX
SQ   Sequence   67 AA;


  Query Match             100.0%;  Score 47;  DB 18;  Length 67;
  Best Local Similarity   100.0%;  Pred. No. 0.69;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db      16 schlpwa 22


RESULT  10
AAW34400
ID   AAW34400 standard; Protein; 77 AA.
XX
AC   AAW34400;
XX
DT   04-MAR-1998  (first entry)
XX
DE   Human Met-OB protein residues 40-116.
XX
KW   Obesity protein; OB; Met-OB; mouse; human; lean tissue mass; liposuction;
KW   athletic performance improvement; implant surgery; insulin sensitivity;
KW   cardiac surgery; bone resorption; diabetes; osteoporosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 11
FT                   /note= "optionally substituted"
FT   Misc-difference 14
FT                   /note= "optionally substituted"
FT   Misc-difference 21
FT                   /note= "optionally substituted"
FT   Misc-difference 25
FT                   /note= "optionally substituted"
FT   Misc-difference 27
FT                   /note= "optionally substituted"
FT   Misc-difference 28
FT                   /note= "optionally substituted"
FT   Misc-difference 29
FT                   /note= "optionally substituted"
FT   Misc-difference 32
FT                   /note= "optionally substituted"
FT   Misc-difference 35
FT                   /note= "optionally substituted"
FT   Misc-difference 38
FT                   /note= "optionally substituted"
FT   Misc-difference 39
FT                   /note= "optionally substituted"
FT   Misc-difference 50
FT                   /note= "optionally substituted"
FT   Misc-difference 58
FT                   /note= "optionally substituted"
FT   Misc-difference 61
FT                   /note= "optionally substituted"
FT   Misc-difference 63
FT                   /note= "optionally substituted"
FT   Misc-difference 66
FT                   /note= "optionally substituted"
FT   Misc-difference 67
FT                   /note= "optionally substituted"
FT   Misc-difference 68
FT                   /note= "optionally substituted"
FT   Misc-difference 69
FT                   /note= "optionally substituted"
FT   Misc-difference 72
FT                   /note= "optionally substituted"
FT   Misc-difference 73
FT                   /note= "optionally substituted"
XX
PN   WO9718833-A1.
XX
PD   29-MAY-1997.
XX
PF   04-NOV-1996;    96WO-US17718.
XX
PR   22-NOV-1995;    95US-0561732.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Pelleymounter MA,  Toombs CF;
XX
DR   WPI; 1997-310245/28.
XX
PT   Increasing lean tissue content by using obesity protein analogue -
PT   for cosmetic use, to improve athletic performance, increase insulin
PT   sensitivity, overall body strength and decrease bone resorption
XX
PS   Claim 1; Page -; 50pp; English.
XX
CC   AAW34394-W34401 represent mutations and truncations of the mouse and
CC   human obesity (OB) proteins shown in AAW22434 and AAW22435. The OB
CC   proteins, and these mutations can all be used in the method of the
CC   invention. The method of the invention is for increasing lean tissue mass
CC   by administration of an OB analogue. The OB proteins are used for
CC   cosmetic applications, to improve athletic performance or as adjunct to
CC   surgery (e.g. liposuction or implant surgery, cardiac surgery, treatment
CC   of broken bones etc.). OB proteins also increase insulin sensitivity and
CC   overall body strength, and decrease bone resorption, e.g. for treating
CC   diabetes or reducing the amount of insulin needed to treat this disease,
CC   or to reverse/improve frailty caused by osteoporosis. Also red blood
CC   cell production is increased, increasing mental as well as physical
CC   performance. OB are already known to reduce weight but are now found to
CC   increase lean mass in non-obese subjects. These OB proteins do not have
CC   the side effects associated with use of anabolic steroids, growth
CC   hormone etc. Chemical modification of the OB protein (e.g. by attachment
```

```
CC   to a water soluble polymer) improves stability, increases circulation
CC   time and/or reduces immunogenicity.
XX
SQ   Sequence   77 AA;


  Query Match           100.0%;  Score 47;  DB 18;  Length 77;
  Best Local Similarity  100.0%;  Pred. No. 0.78;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          |||||||
Db     56 schlpwa 62


RESULT  11
AAW27173
ID   AAW27173 standard; Protein; 77 AA.
XX
AC   AAW27173;
XX
DT   09-DEC-1997  (first entry)
XX
DE   Human recombinant truncated OB protein analogue.
XX
KW   Obesity; blood lipid; weight loss; cholesterol; arterial plaque;
KW   familial hypercholesterolaemia; triglyceride; hypertension;
KW   gall stone formation.
XX
OS   Homo sapiens.
XX
PN   WO9706816-A1.
XX
PD   27-FEB-1997.
XX
PF   02-AUG-1996;   96WO-US12674.
XX
PR   17-AUG-1995;   95US-0516263.
XX
PA   (AMGE-) AMGEN INC.
XX
PI   Pelleymounter MA;
XX
DR   WPI; 1997-178778/16.
XX
PT   Reducing levels of blood lipid(s) without inducing weight loss - by
PT   administering human or murine mature OB protein or their specified
PT   derivatives
XX
PS   Claim 1; Page -; 52pp; English.
XX
CC   A method has been developed for reducing the level of blood lipids (BL)
CC   in non-obese patients (NOP), or maintaining reduced levels of BL in NOP
CC   having an elevated level of BL. The method involves administering an OB
CC   protein, analogue or derivative, in an amount insufficient to cause
CC   weight loss. The present sequence represents a recombinant truncated
CC   human OB protein analogue from the wild-type positions 40-116, which can
CC   be used in the above method. The method is used to treat conditions
CC   related to BL levels, including high cholesterol, particularly familial
CC   hypercholesterolaemia, high triglyceride levels, arterial plaque and
CC   hypertension, and to prevent gall stone formation. The method does not
CC   result in weight reduction or further weight loss.
CC   N.B. The present sequence is not shown in the specification, but is
CC   derived from SEQ ID NO:4 as specified in claim 1, where the positions
CC   are numbered without the N-terminal methionine residue.
XX
SQ   Sequence   77 AA;


  Query Match           100.0%;  Score 47;  DB 18;  Length 77;
  Best Local Similarity  100.0%;  Pred. No. 0.78;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          |||||||
Db     56 schlpwa 62


RESULT  12
AAR98746
ID   AAR98746 standard; protein; 105 AA.
XX
AC   AAR98746;
XX
DT   14-NOV-1996  (first entry)
XX
DE   Biologically active anti-obesity protein.
XX
KW   anti-obesity; regulate; fat tissue; obesity; type II diabetes;
KW   cardiovascular disease; cancer.
XX
OS   Synthetic.
XX
PN   US5532336-A.
XX
PD   02-JUL-1996.
XX
PF   31-JAN-1995;   95US-0381034.
XX
PR   31-JAN-1995;   95US-0381034.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Dimarchi RD,  Flora DB,  Heath WF,  Hoffmann JA,  Shields JE;
PI   Smiley DL;
XX
DR   WPI; 1996-321178/32.
XX
PT   Anti-obesity peptide - reduces the risk of, e.g. type II diabetes or
PT   cancer, in obese individuals
XX
PS   Claim 17; Column -; 11pp; English.
XX
CC   The present sequence is a biologically active anti-obesity protein. When
CC   administered to a patient the protein regulates fat tissue, allowing
CC   patients to overcome their obesity handicap and live normal lives with
CC   much reduced risk for type II diabetes, cardiovascular disease and
CC   cancer. A specifically claimed protein is shown in AAR98746.
XX
SQ   Sequence   105 AA;


  Query Match           100.0%;  Score 47;  DB 17;  Length 105;
  Best Local Similarity  100.0%;  Pred. No. 1;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          |||||||
Db     54 schlpwa 60


RESULT  13
AAW02150
ID   AAW02150 standard; Protein; 119 AA.
XX
AC   AAW02150;
XX
DT   10-NOV-1996  (first entry)
XX
DE   Anti-obesity protein.
XX
KW   Anti-obesity protein; fat; weight control; food additive.
XX
OS   Synthetic.
XX
PN   WO9623518-A1.
```

XX
PD 08-AUG-1996.
XX
PF 29-JAN-1996; 96WO-US01345.
XX
PR 06-FEB-1995; 95US-0383639.
PR 31-JAN-1995; 95US-0381031.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Basinski MB, Dimarchi RD, Heath WF, Schoner BE;
XX
DR WPI; 1996-371128/37.
XX
PT Protein for treatment of obesity and obesity related conditions -
PT also useful as feed additive, to raise antibodies for diagnostic use
PT and to control wt. in mammals, i.e. to improve bodily appearance
XX
PS Claim 15; Page 28; 35pp; English.
XX
CC An anti-obesity protein (AAW02150) is a specific example of the
CC generic anti-obesity protein formula given in AAW02135. This
CC anti-obesity protein and others (see also AAW02136-49, AAW00861) can
CC be obtd. by chemical synthesis or semi-synthesis, or by recombinant
CC DNA technology and expression in host cells. They provide effective
CC treatment for obesity and many offer additional advantages of better
CC adsorption characteristics and improved in vivo stability.
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 47; DB 17; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 95 schlpwa 101

RESULT 14
AAR92732
ID AAR92732 standard; Protein; 119 AA.
XX
AC AAR92732;
XX
DT 13-SEP-1996 (first entry)
XX
DE Obesity protein C-terminal fragment.
XX
KW Obesity; mouse; OBP; leptin; hormone; body weight regulation; diabetes;
KW food intake; energy expenditure; high blood pressure; cholesterol; human;
KW gene therapy; antibody; cancer; Kobe beef; Foie gras; immunoassay.
XX
OS Homo sapiens.
XX
PN GB2292382-A.
XX
PD 21-FEB-1996.
XX
PF 17-AUG-1995; 95GB-0016947.
XX
PR 07-JUN-1995; 95US-0483211.
PR 17-AUG-1994; 94US-0292345.
PR 30-NOV-1994; 94US-0347563.
PR 10-MAY-1995; 95US-0438431.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Burley SK, Friedman JM, Gajiwala K, Halaas JL, Maffei M;
PI Proenca R, Zhang Y;
XX
DR WPI; 1996-099009/11.
DR N-PSDB; AAT16375.
XX
PT Obesity polypeptide(s) able to modulate body wt. - useful for e.g.
PT reducing wt. in treatment of diabetes, high blood pressure and high
PT cholesterol and for cosmetic reasons
XX
PS Claim 27; Page 184; 304pp; English.
XX
CC AAR92731 and AAR92732 represent fragments of the human obesity
CC polypeptide (OBP) (see AAR92720 for full length sequence). This sequence
CC represents the C-terminus of OBP. OBP (also known as leptin) is a
CC hormone involved in the regulation of body weight. The full length
CC protein sequence has effects on both food intake and energy expenditure.
CC OBP and its analogues are useful for modifying body weight (optionally
CC combined with known medicaments), for treating diabetes, high blood
CC pressure or high cholesterol. The full length OBP coding sequence (and
CC sequences complimentary to it) can be used in gene therapy for modifying
CC body weight. The full length protein can be used for reducing weight
CC for health or cosmetic reasons in obese humans, or to produce leaner
CC food animals. Antagonists of OBP (including antibodies) are useful for
CC increasing body weight, e.g. for treating weight loss associated with
CC cancer, or for cosmetic reasons in humans, or for production of Kobe
CC beef or Foie gras in domestic animals. OBP antibodies (Ab) can also be
CC used in diagnostic immunoassays for the presence of OBP. The formation
CC of Ab-OBP complexes enables in vitro evaluation of levels of OBP in a
CC sample, especially to detect diseases associated with elevated or
CC decreased levels, and to monitor treatment of these diseases.
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 47; DB 17; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 68 schlpwa 74

RESULT 15
AAB28458
ID AAB28458 standard; Protein; 119 AA.
XX
AC AAB28458;
XX
DT 01-FEB-2001 (first entry)
XX
DE Human OB protein C-terminal portion.
XX
KW Human; mouse; OB gene; obesity; adiposity; body weight.
XX
OS Homo sapiens.
XX
PN US6124448-A.
XX
PD 26-SEP-2000.
XX
PF 07-JUN-1995; 95US-0488208.
XX
PR 17-AUG-1994; 94US-0292345.
PR 30-NOV-1994; 94US-0347563.
PR 10-MAY-1995; 95US-0438431.
XX
PA (UYRQ) UNIV ROCKEFELLER.
XX
PI Maffei M, Proenca R, Zhang Y, Friedman JM;
XX
DR WPI; 2000-601556/57.
DR N-PSDB; AAC62578.
XX
PT Nucleic acid primers and probes useful for detecting mutations in
PT mammalian OB gene associated with regulation of body weight and

PT adiposity -
XX
PS Disclosure; Column 123-124; 153pp; English.
XX
CC The present sequence is encoded by a nucleotide sequence used in an
CC invention relating to the control of body weight of animals including
CC humans. Nucleic acids of at least 10 nucleotides which are hybridisable
CC to a non-coding region of an OB nucleic acid have been created. The OB
CC gene plays a critical role in the regulation of body weight and
CC adiposity. The nucleic acids may be used as probes or as primers for PCR.
CC They are useful for evaluating the presence of mutations in the human OB
CC gene or for evaluating the level of expression of OB mRNA. Defects
CC associated with OB gene expression result in obese phenotypes.
XX
SQ Sequence 119 AA;

Query Match 100.0%; Score 47; DB 21; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 68 schlpwa 74

Search completed: July 3, 2001, 20:47:13
Job time: 38 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 3, 2001, 20:47:41 ; Search time 9.7 Seconds
(without alignments)
24.720 Million cell updates/sec

Title: US-09-377-081-18
Perfect score: 47
Sequence: 1 SCHLPWA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed; and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	47	100.0	146	1	OB_GORGO	Q95189 gorilla gor
2	47	100.0	146	1	OB_PANTR	O02750 pan troglod
3	47	100.0	146	1	OB_PONPY	Q95234 pongo pygma
4	47	100.0	167	1	OB_HUMAN	P41159 homo sapien
5	37	78.7	1401	1	WRN_MOUSE	O09053 mus musculu
6	35	74.5	409	1	MDP1_PIG	P22412 sus scrofa
7	35	74.5	410	1	MDP1_SHEEP	P43477 ovis aries
8	35	74.5	411	1	MDP1_HUMAN	P16444 homo sapien
9	34	72.3	167	1	OB_MACMU	Q28504 macaca mula
10	34	72.3	376	1	CGD3_ARATH	P42753 arabidopsis
11	34	72.3	421	1	ACRO_HUMAN	P10323 homo sapien
12	34	72.3	453	1	PABB_ECOLI	P05041 escherichia
13	34	72.3	454	1	PABB_SALTY	P12680 salmonella
14	34	72.3	1432	1	WRN_HUMAN	Q14191 homo sapien
15	33	70.2	287	1	YJJW_ECOLI	P39409 escherichia
16	33	70.2	387	1	GSPL_ECOLI	P45763 escherichia
17	33	70.2	531	1	UD13_RAT	Q64637 rattus norv
18	33	70.2	533	1	UD12_MOUSE	P70691 mus musculu
19	33	70.2	533	1	UD12_RAT	P20720 rattus norv
20	32.5	69.1	1630	1	ESP1_YEAST	Q03018 saccharomyc
21	32	68.1	188	1	NRFB_ECOLI	P32707 escherichia
22	32	68.1	410	1	MDP1_MOUSE	P31428 mus musculu
23	32	68.1	410	1	MDP1_RABIT	P31429 oryctolagus
24	32	68.1	410	1	MDP1_RAT	P31430 rattus norv
25	32	68.1	464	1	GRFR_RAT	Q02644 rattus norv
26	32	68.1	474	1	CBLC_HUMAN	Q9ulv8 homo sapien
27	32	68.1	493	1	HXKL_ARATH	Q9t071 arabidopsis
28	32	68.1	532	1	UD14_RABIT	Q28612 oryctolagus
29	32	68.1	552	1	DAK_CITFR	P45510 citrobacter
30	32	68.1	621	1	FOR_THELI	Q56303 thermococcu
31	32	68.1	650	1	BM86_BOOMI	P20736 boophilus m
32	32	68.1	805	1	YGW6_YEAST	P53086 saccharomyc
33	32	68.1	825	1	IL4R_HUMAN	P24394 homo sapien
34	32	68.1	1357	1	POL2_TBRVS	P14547 tomato blac
35	32	68.1	1411	1	YK63_CAEEL	P34342 caenorhabdi
36	32	68.1	1480	1	SLIT_DROME	P24014 drosophila
37	32	68.1	2431	1	POLN_SFV	P08411 semliki for
38	32	68.1	2514	1	POLN_ONNVG	P13886 o'nyong-nyo
39	32	68.1	3206	1	POLG_PSBMV	P29152 p genome po
40	31.5	67.0	353	1	RHL_HYLPI	Q28446 hylobates p
41	31.5	67.0	416	1	RHCE_HUMAN	P18577 homo sapien
42	31.5	67.0	416	1	RHD_HUMAN	Q02161 homo sapien
43	31.5	67.0	416	1	RHLA_PANTR	Q28813 pan troglod
44	31.5	67.0	416	1	RHLC_GORGO	Q28426 gorilla gor
45	31.5	67.0	416	1	RHLD_GORGO	Q28427 gorilla gor

ALIGNMENTS

```
RESULT   1
OB_GORGO
ID   OB_GORGO       STANDARD;      PRT;   146 AA.
AC   Q95189;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LEPTIN (OBESITY FACTOR).
GN   LEP OR OB.
OS   Gorilla gorilla gorilla (Lowland gorilla).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX   NCBI_TaxID=9595;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.P., Zhang X., Hsiung H.M.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS
CC       TO REGULATE THE SIZE OF THE BODY FAT DEPOT. AN INCREASE IN THE
CC       LEVEL OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT
CC       FOOD INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A
CC       HOMEOSTATIC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS.
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U72872; AAB17091.1; -.
DR   InterPro; IPR000065; -.
DR   Pfam; PF02024; Leptin; 1.
KW   Obesity.
FT   DISULFID     96    146       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  16031 MW;  02C43BF6B9A4C85C CRC64;


  Query Match          100.0%;  Score 47;  DB 1;  Length 146;
  Best Local Similarity  100.0%;  Pred. No. 0.19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db      95 SCHLPWA 101


RESULT   2
OB_PANTR
ID   OB_PANTR       STANDARD;      PRT;   146 AA.
AC   O02750;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
```

```
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LEPTIN (OBESITY FACTOR).
GN   LEP OR OB.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schoner B., Basinski M.B., Smith D.P., Hsiung H.M., Zhang X.,
RA   Rockey P.K., Rosteck P.R.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS
CC       TO REGULATE THE SIZE OF THE BODY FAT DEPOT. AN INCREASE IN THE
CC       LEVEL OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT
CC       FOOD INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A
CC       HOMEOSTATIC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS.
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96450; AAB54023.1; -.
DR   InterPro; IPR000065; -.
DR   Pfam; PF02024; Leptin; 1.
KW   Obesity.
FT   DISULFID     96    146       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  16058 MW;  02C42A06B554D55C CRC64;


  Query Match             100.0%;  Score 47;  DB 1;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db      95 SCHLPWA 101


RESULT   3
OB_PONPY
ID   OB_PONPY       STANDARD;      PRT;   146 AA.
AC   Q95234;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LEPTIN (OBESITY FACTOR).
GN   LEP OR OB.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Smith D.P., Zhang X., Hsiung H.M.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS
CC       TO REGULATE THE SIZE OF THE BODY FAT DEPOT. AN INCREASE IN THE
CC       LEVEL OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT
CC       FOOD INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A
CC       HOMEOSTATIC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS.
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U72873; AAB17092.1; -.
DR   InterPro; IPR000065; -.
DR   Pfam; PF02024; Leptin; 1.
KW   Obesity.
FT   DISULFID     96    146       BY SIMILARITY.
SQ   SEQUENCE   146 AA;  16195 MW;  3F50A13338FFDBD4 CRC64;


  Query Match             100.0%;  Score 47;  DB 1;  Length 146;
  Best Local Similarity   100.0%;  Pred. No. 0.19;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db      95 SCHLPWA 101


RESULT   4
OB_HUMAN
ID   OB_HUMAN       STANDARD;      PRT;   167 AA.
AC   P41159; O15158;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   LEPTIN PRECURSOR (OBESITY FACTOR) (OBESE PROTEIN).
GN   LEP OR OB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95075453; PubMed=7984236;
RA   Zhang Y., Proenca P., Maffei M., Barone M., Leopold L.,
RA   Friedman J.M.;
RT   "Positional cloning of the mouse obese gene and its human homologue.";
RL   Nature 372:425-432(1994).
RN   [2]
RP   ERRATUM.
RA   Zhang Y., Proenca P., Maffei M., Barone M., Leopold L.,
RA   Friedman J.M.;
RL   Nature 374:479-479(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95309556; PubMed=7789654;
RA   Masuzaki H., Ogawa Y., Isse N., Satoh N., Okazaki T.,
RA   Shigemoto M., Mori K., Tamura N., Hosoda K., Yoshimasa Y.,
RA   Jingami H., Kawada T., Nakao K.;
RT   "Human obese gene expression. Adipocyte-specific expression and
RT   regional differences in the adipose tissue.";
RL   Diabetes 44:855-858(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96223958; PubMed=8626726;
RA   Gong D.W., Bi S., Pratley R.E., Weintraub B.D.;
RT   "Genomic structure and promoter analysis of the human obese gene.";
RL   J. Biol. Chem. 271:3971-3974(1996).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Chehab F.F., Lim M.E.;
RL   Submitted (DEC-1995) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96070903; PubMed=7499240;
RA   Isse N., Ogawa Y., Tamura N., Masuzaki H., Mori K., Okazaki T.,
RA   Satoh N., Shigemoto M., Yoshimasa Y., Nishi S., Hosada K., Inazawa J.,
RA   Nakao K.;
RT   "Structural organization and chromosomal assignment of the human
```

```
RT   obese gene.";
RL   J. Biol. Chem. 270:27728-27733(1995).
RN   [7]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96198511; PubMed=8621021;
RA   Niki T., Mori H., Tamori Y., Kishimoto-Hashiramoto M., Ueno H.,
RA   Araki S., Masugi J., Sawant N., Majithia H.R., Rais N.,
RA   Hashiramoto M., Taniguchi H., Kasuga M.;
RT   "Human obese gene: molecular screening in Japanese and Asian Indian
RT   NIDDM patients associated with obesity.";
RL   Diabetes 45:675-678(1996).
RN   [8]
RP   SEQUENCE FROM N.A.
RA   Lu L., Fu Z., Xu M., Fu Y., Hu Z.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [9]
RP   STRUCTURE BY NMR.
RX   MEDLINE=97309492; PubMed=9166907;
RA   Kline A.D., Becker G.W., Churgay L.M., Landen B.E., Martin D.K.,
RA   Muth W.L., Rathnachalam R., Richardson J.M., Schoner B., Ulmer M.,
RA   Hale J.E.;
RT   "Leptin is a four-helix bundle: secondary structure by NMR.";
RL   FEBS Lett. 407:239-242(1997).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX   MEDLINE=97289390; PubMed=9144295;
RA   Zhang F., Basinski M.B., Beals J.M., Briggs S.L., Churgay L.M.,
RA   Clawson D.K., Dimarchi R.D., Furman T.C., Hale J.E., Hsiung H.M.,
RA   Schoner B.E., Smith D.P., Zhang X.Y., Wery J.P., Schevitz R.W.;
RT   "Crystal structure of the obese protein leptin-E100.";
RL   Nature 387:206-209(1997).
RN   [11]
RP   VARIANT MET-94.
RA   Bartholomew D.W., McClellan J.M.;
RT   "A novel polymorphism in the leptin gene.";
RL   Hum. Mutat. 12:220-220(1998).
RN   [12]
RP   VARIANT TRP-105.
RX   MEDLINE=98160176; PubMed=9500540;
RA   Strobel A., Issad T., Camoin L., Ozata M., Strosberg A.D.;
RT   "A leptin missense mutation associated with hypogonadism and morbid
RT   obesity.";
RL   Nat. Genet. 18:213-215(1998).
CC   -!- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS
CC       TO REGULATE THE SIZE OF THE BODY FAT DEPOT. AN INCREASE IN THE
CC       LEVEL OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT
CC       FOOD INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A
CC       HOMEOSTATIC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS.
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18915; AAA60470.1; -.
DR   EMBL; D49487; BAA08448.1; -.
DR   EMBL; U43653; AAC50400.1; -.
DR   EMBL; U43415; AAC31660.1; -.
DR   EMBL; D63710; BAA09839.1; -.
DR   EMBL; D63709; BAA09839.1; JOINED.
DR   EMBL; D63519; BAA09787.1; -.
DR   EMBL; D63518; BAA09787.1; JOINED.
DR   EMBL; AF008123; AAB63507.1; -.
DR   PDB; 1AX8; 13-JAN-99.
DR   MIM; 164160; -.
DR   InterPro; IPR000065; -.
DR   Pfam; PF02024; Leptin; 1.
DR   PRINTS; PR00495; LEPTIN.
KW   Diabetes; Obesity; Signal; Polymorphism; Disease mutation;
KW   3D-structure.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    167       LEPTIN.
FT   DISULFID   117    167
FT   VARIANT     49     49       MISSING (IN 30% THE CLONES).
FT                               /FTId=VAR_004196.
FT   VARIANT     94     94       V -> M.
FT                               /FTId=VAR_004197.
FT   VARIANT    105    105       R -> W (IN MORBID OBESITY AND
FT                               HYPOGONADISM).
FT                               /FTId=VAR_008094.
FT   CONFLICT    96     96       Q -> R (IN REF. 8).
SQ   SEQUENCE   167 AA;  18640 MW;  C91A121E92D37B69 CRC64;


  Query Match          100.0%;  Score 47;  DB 1;  Length 167;
  Best Local Similarity  100.0%;  Pred. No. 0.21;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||||||
Db     116 SCHLPWA 122


RESULT   5
WRN_MOUSE
ID   WRN_MOUSE      STANDARD;      PRT;  1401 AA.
AC   O09053; O09050; Q9Z242;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WERNER SYNDROME HELICASE HOMOLOG.
GN   WRN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C; TISSUE=Testis, and Spleen;
RX   MEDLINE=97288537; PubMed=9143515;
RA   Imamura O., Ichikawa K., Yamabe Y., Goto M., Sugawara M.,
RA   Furuichi Y.;
RT   "Cloning of a mouse homologue of the human Werner syndrome gene and
RT   assignment to 8A4 by fluorescence in situ hybridization.";
RL   Genomics 41:298-300(1997).
RN   [2]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=98284027; PubMed=9618508;
RA   Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT   "Nucleolar localization of the Werner syndrome protein in human
RT   cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Paeper B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA   Mulligan J., Galas D., Fu Y.-H.;
RT   "Genomic structure of the human Werner's gene and cloning of its mouse
RT   homolog.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF GENOMIC STABILITY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: BELONGS TO THE RECQ SUBFAMILY OF HELICASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; D86527; BAA20270.1; -.
DR   EMBL; D86526; BAA20269.1; -.
DR   EMBL; AF091215; AAC78077.1; -.
DR   MGD; MGI:109635; Wrn.
DR   InterPro; IPR001410; -.
DR   InterPro; IPR001650; -.
DR   InterPro; IPR002121; -.
DR   InterPro; IPR002562; -.
DR   Pfam; PF01612; 3_5_exonuclease; 1.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00570; HRDC; 1.
DR   Pfam; PF00271; helicase_C; 1.
KW   Helicase; ATP-binding; Nuclear protein.
FT   NP_BIND     535    542       ATP (BY SIMILARITY).
FT   SITE        632    635       DEAH BOX.
FT   DOMAIN     1115   1194       HRDC.
FT   DOMAIN     1387   1390       POLY-SER.
FT   CONFLICT    101    101       N -> S (IN REF. 3).
FT   CONFLICT    228    228       V -> A (IN REF. 3).
FT   CONFLICT    250    250       L -> S (IN REF. 3).
FT   CONFLICT    452    452       M -> V (IN REF. 3).
FT   CONFLICT    459    459       K -> T (IN REF. 3).
FT   CONFLICT    468    468       C -> R (IN REF. 3).
FT   CONFLICT    619    619       K -> Q (IN REF. 3).
FT   CONFLICT    800    800       Q -> K (IN REF. 3).
FT   CONFLICT   1021   1021       L -> S (IN REF. 3).
FT   CONFLICT   1145   1145       A -> T (IN REF. 3).
FT   CONFLICT   1181   1182       VG -> LE (IN REF. 3).
FT   CONFLICT   1252   1252       V -> A (IN REF. 3).
FT   CONFLICT   1308   1308       I -> L (IN REF. 3).
FT   CONFLICT   1356   1356       V -> A (IN REF. 3).
SQ   SEQUENCE   1401 AA;  157256 MW;  949060992467FB8C CRC64;
```

```
  Query Match             78.7%;  Score 37;  DB 1;  Length 1401;
  Best Local Similarity   85.7%;  Pred. No. 56;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           |||| ||
Db     828 SCHLLWA 834
```

```
RESULT   6
MDP1_PIG
ID   MDP1_PIG       STANDARD;      PRT;   409 AA.
AC   P22412;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MICROSOMAL DIPEPTIDASE PRECURSOR (EC 3.4.13.19) (MDP)
DE   (DEHYDROPEPTIDASE-I) (RENAL DIPEPTIDASE) (RDP).
GN   DPEP1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Kidney cortex;
RX   MEDLINE=91058511; PubMed=2173907;
RA   Rached E., Hooper N.M., James P., Semenza G., Turner A.J.,
RA   Mantei N.;
RT   "cDNA cloning and expression in Xenopus laevis oocytes of pig renal
RT   dipeptidase, a glycosyl-phosphatidylinositol-anchored ectoenzyme.";
RL   Biochem. J. 271:755-760(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Satoh S., Koyama S., Ohtuka K., Keida Y., Niwa M., Kohsaka M.;
RL   Submitted (SEP-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 17-39.
RX   MEDLINE=90147607; PubMed=2137335;
RA   Hooper N.M., Keen J.N., Turner A.J.;
RT   "Characterization of the glycosyl-phosphatidylinositol-anchored human
RT   renal dipeptidase reveals that it is more extensively glycosylated
RT   than the pig enzyme.";
RL   Biochem. J. 265:429-433(1990).
CC   -!- FUNCTION: HYDROLYZES A WIDE RANGE OF DIPEPTIDES. IMPLICATED IN THE
CC       RENAL METABOLISM OF GLUTATHIONE AND ITS CONJUGATES. CONVERTS
CC       LEUKOTRIENE D4 TO LEUKOTRIENE E4; IT MAY PLAY AN IMPORTANT ROLE IN
CC       THE REGULATION OF LEUKOTRIENE ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O = 2 AMINO ACID.
CC   -!- COFACTOR: ZINC.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR;
CC       BRUSH BORDER MEMBRANE.
CC   -!- PTM: THE PRECISE POSITION OF THE C-TERMINUS AND GPI-ANCHOR OF THE
CC       MATURE RENAL DIPEPTIDASE IS NOT YET KNOWN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M19.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X53730; CAA37762.1; -.
DR   EMBL; D13142; BAA02433.1; -.
DR   PIR; JS0759; JS0759.
DR   PIR; PS0394; PS0394.
DR   PIR; S08194; S08194.
DR   PIR; S13059; S13059.
DR   MEROPS; M19.001; -.
DR   InterPro; IPR000180; -.
DR   Pfam; PF01244; Renal_dipeptase; 1.
DR   PROSITE; PS00869; RENAL_DIPEPTIDASE; 1.
KW   Hydrolase; Dipeptidase; Metalloprotease; Zinc; Microsome; Signal;
KW   GPI-anchor; Glycoprotein.
FT   SIGNAL        1     16
FT   CHAIN        17    384       MICROSOMAL DIPEPTIDASE.
FT   PROPEP      385    409       REMOVED IN MATURE FORM (BY SIMILARITY).
FT   ACT_SITE    141    141       BY SIMILARITY.
FT   METAL       286    286       ZINC (CATALYTIC) (POTENTIAL).
FT   METAL       289    289       ZINC (CATALYTIC) (POTENTIAL).
FT   CARBOHYD     57     57       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   CARBOHYD    279    279       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID       384    384       GPI-ANCHOR (BY SIMILARITY).
SQ   SEQUENCE   409 AA;  44700 MW;  926B7F0044FA055F CRC64;
```

```
  Query Match             74.5%;  Score 35;  DB 1;  Length 409;
  Best Local Similarity   71.4%;  Pred. No. 43;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           ||: |||
Db     169 SCNTPWA 175
```

```
RESULT   7
MDP1_SHEEP
ID   MDP1_SHEEP     STANDARD;      PRT;   410 AA.
AC   P43477;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MICROSOMAL DIPEPTIDASE PRECURSOR (EC 3.4.13.19) (MDP)
DE   (DEHYDROPEPTIDASE-I) (RENAL DIPEPTIDASE) (RDP).
GN   DPEP1.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
```

```
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 17-56.
RC   TISSUE=Lung;
RX   MEDLINE=94331427; PubMed=8054366;
RA   An S., Schmidt F.J., Campbell B.J.;
RT   "Molecular cloning of sheep lung dipeptidase: a glycosyl
RT   phosphatidylinositol-anchored ectoenzyme that converts leukotriene D4
RT   to leukotriene E4.";
RL   Biochim. Biophys. Acta 1226:337-340(1994).
CC   -!- FUNCTION: HYDROLYZES A WIDE RANGE OF DIPEPTIDES. IMPLICATED IN THE
CC       RENAL METABOLISM OF GLUTATHIONE AND ITS CONJUGATES. CONVERTS
CC       LEUKOTRIENE D4 TO LEUKOTRIENE E4; IT MAY PLAY AN IMPORTANT ROLE IN
CC       THE REGULATION OF LEUKOTRIENE ACTIVITY. IN LUNG TISSUE, IT MAY
CC       TERMINATE OR SIGNIFICANTLY REDUCE THE LEUKOTRIENE INDUCED SIGNAL
CC       FOR BRONCHOSPASM.
CC   -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O = 2 AMINO ACID.
CC   -!- COFACTOR: ZINC.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR;
CC       BRUSH BORDER MEMBRANE.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN LUNG, KIDNEY AND INTESTINAL
CC       TISSUES.
CC   -!- PTM: THE PRECISE POSITION OF THE C-TERMINUS AND GPI-ANCHOR OF THE
CC       MATURE RENAL DIPEPTIDASE IS NOT YET KNOWN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M19.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L27113; AAA21725.1; -.
DR   MEROPS; M19.001; -.
DR   InterPro; IPR000180; -.
DR   Pfam; PF01244; Renal_dipeptase; 1.
DR   PROSITE; PS00869; RENAL_DIPEPTIDASE; 1.
KW   Hydrolase; Dipeptidase; Metalloprotease; Zinc; Microsome; Signal;
KW   GPI-anchor; Glycoprotein.
FT   SIGNAL       1    16       BY SIMILARITY.
FT   CHAIN       17   384       MICROSOMAL DIPEPTIDASE.
FT   PROPEP     385   410       REMOVED IN MATURE FORM (BY SIMILARITY).
FT   ACT_SITE   141   141       BY SIMILARITY.
FT   METAL      286   286       ZINC (CATALYTIC) (POTENTIAL).
FT   LIPID      384   384       GPI-ANCHOR (BY SIMILARITY).
FT   CARBOHYD    57    57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    62    62       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   279   279       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    42    42       A -> Q (IN AA SEQUENCE).
SQ   SEQUENCE   410 AA;  45096 MW;  AA818C8B8BB91F31 CRC64;


  Query Match          74.5%;  Score 35; DB 1;  Length 410;
  Best Local Similarity  71.4%;  Pred. No. 43;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           ||: |||
Db     169 SCNTPWA 175


RESULT   8
MDP1_HUMAN
ID   MDP1_HUMAN     STANDARD;      PRT;   411 AA.
AC   P16444;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   MICROSOMAL DIPEPTIDASE PRECURSOR (EC 3.4.13.19) (MDP)
DE   (DEHYDROPEPTIDASE-I) (RENAL DIPEPTIDASE) (RDP).
GN   DPEP1 OR RDP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93176806; PubMed=8439558;
RA   Satoh S., Kusunoki C., Konta Y., Niwa M., Kohsaka M.;
RT   "Cloning and structural analysis of genomic DNA for human renal
RT   dipeptidase.";
RL   Biochim. Biophys. Acta 1172:181-183(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90154088; PubMed=2303490;
RA   Adachi H., Tawaragi Y., Inuzuka C., Kubota I., Tsujimoto M.,
RA   Nishihara T., Nakazato H.;
RT   "Primary structure of human microsomal dipeptidase deduced from
RT   molecular cloning.";
RL   J. Biol. Chem. 265:3992-3995(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=94226762; PubMed=7764673;
RA   Satoh S., Ohtsuka K., Keida Y., Kusunoki C., Konta Y., Niwa M.,
RA   Kohsaka M.;
RT   "Gene structural analysis and expression of human renal dipeptidase.";
RL   Biotechnol. Prog. 10:134-140(1994).
RN   [4]
RP   SEQUENCE OF 17-56; 111-121 AND 298-310.
RX   MEDLINE=90147607; PubMed=2137335;
RA   Hooper N.M., Keen J.N., Turner A.J.;
RT   "Characterization of the glycosyl-phosphatidylinositol-anchored human
RT   renal dipeptidase reveals that it is more extensively glycosylated
RT   than the pig enzyme.";
RL   Biochem. J. 265:429-433(1990).
RN   [5]
RP   SEQUENCE OF 17-39.
RX   MEDLINE=89359222; PubMed=2768222;
RA   Adachi H., Kubota I., Okamura N., Iwata H., Tsujimoto M., Nakazato H.,
RA   Nishihara T., Noguchi T.;
RT   "Purification and characterization of human microsomal dipeptidase.";
RL   J. Biochem. 105:957-961(1989).
RN   [6]
RP   GPI-ANCHOR.
RX   MEDLINE=90368722; PubMed=2168407;
RA   Adachi H., Katayama T., Inuzuka C., Oikawa S., Tsujimoto M.,
RA   Nakazato H.;
RT   "Identification of membrane anchoring site of human renal dipeptidase
RT   and construction and expression of a cDNA for its secretory form.";
RL   J. Biol. Chem. 265:15341-15345(1990).
RN   [7]
RP   ACTIVE SITE.
RX   MEDLINE=93237320; PubMed=8097406;
RA   Adachi H., Katayama T., Nakazato H., Tsujimoto M.;
RT   "Importance of Glu-125 in the catalytic activity of human renal
RT   dipeptidase.";
RL   Biochim. Biophys. Acta 1163:42-48(1993).
CC   -!- FUNCTION: HYDROLYZES A WIDE RANGE OF DIPEPTIDES. IMPLICATED IN THE
CC       RENAL METABOLISM OF GLUTATHIONE AND ITS CONJUGATES. CONVERTS
CC       LEUKOTRIENE D4 TO LEUKOTRIENE E4; IT MAY PLAY AN IMPORTANT ROLE IN
CC       THE REGULATION OF LEUKOTRIENE ACTIVITY.
CC   -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)O = 2 AMINO ACID.
CC   -!- COFACTOR: ZINC.
CC   -!- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR;
CC       BRUSH BORDER MEMBRANE.
CC   -!- PTM: THE PRECISE POSITION OF THE C-TERMINUS AND GPI-ANCHOR OF THE
CC       MATURE RENAL DIPEPTIDASE IS NOT YET KNOWN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M19.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; D13137; BAA02430.1; -.
DR   EMBL; D13128; BAA02430.1; JOINED.
DR   EMBL; D13129; BAA02430.1; JOINED.
DR   EMBL; D13130; BAA02430.1; JOINED.
DR   EMBL; D13131; BAA02430.1; JOINED.
DR   EMBL; D13132; BAA02430.1; JOINED.
DR   EMBL; D13133; BAA02430.1; JOINED.
DR   EMBL; D13134; BAA02430.1; JOINED.
DR   EMBL; D13135; BAA02430.1; JOINED.
DR   EMBL; D13136; BAA02430.1; JOINED.
DR   EMBL; J05257; AAB59410.1; -.
DR   EMBL; D13138; BAA02431.1; -.
DR   EMBL; S70330; AAC60630.2; -.
DR   EMBL; S70329; AAC60630.2; JOINED.
DR   PIR; A35467; A35467.
DR   PIR; JS0756; JS0756.
DR   PIR; JS0757; JS0757.
DR   PIR; PX0021; PX0021.
DR   PIR; S08193; S08193.
DR   PIR; S29848; S29848.
DR   MEROPS; M19.001; -.
DR   MIM; 179780; -.
DR   InterPro; IPR000180; -.
DR   Pfam; PF01244; Renal_dipeptase; 1.
DR   PROSITE; PS00869; RENAL_DIPEPTIDASE; 1.
KW   Hydrolase; Dipeptidase; Metalloprotease; Zinc; Microsome; Signal;
KW   GPI-anchor; Glycoprotein.
FT   SIGNAL       1     16
FT   CHAIN       17    385       MICROSOMAL DIPEPTIDASE.
FT   PROPEP     386    411       REMOVED IN MATURE FORM.
FT   ACT_SITE   141    141
FT   METAL      286    286       ZINC (CATALYTIC) (POTENTIAL).
FT   METAL      289    289       ZINC (CATALYTIC) (POTENTIAL).
FT   CARBOHYD    57     57       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   279    279       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   332    332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   358    358       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   LIPID      385    385       GPI-ANCHOR.
FT   MUTAGEN    141    141       E->Q: PARTIAL LOSS OF ACTIVITY.
FT   MUTAGEN    141    141       E->D,C: COMPLETE LOSS OF ACTIVITY.
FT   CONFLICT     9      9       S -> P (IN REF. 2).
FT   CONFLICT   102    102       M -> R (IN REF. 2).
FT   CONFLICT   125    125       I -> R (IN REF. 2).
SQ   SEQUENCE   411 AA;  45663 MW;  35F7A7F076CB719D CRC64;


  Query Match             74.5%;  Score 35;  DB 1;  Length 411;
  Best Local Similarity   71.4%;  Pred. No. 43;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          ||: ||| |
Db    169 SCNTPWA 175


RESULT   9
OB_MACMU
ID   OB_MACMU       STANDARD;      PRT;   167 AA.
AC   Q28504;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   LEPTIN PRECURSOR (OBESITY FACTOR).
GN   LEP OR OB.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adipose tissue;
RX   MEDLINE=96411743; PubMed=8810296;
RA   Hotta K., Gustafson T.A., Ortmeyer H.K., Bodkin N.L.,
RA   Nicolson M.A., Hansen B.C.;
RT   "Regulation of obese (ob) mRNA and plasma leptin levels in rhesus
RT   monkeys. Effects of insulin, body weight, and non-insulin-dependent
RT   diabetes mellitus.";
RL   J. Biol. Chem. 271:25327-25331(1996).
CC   -!- FUNCTION: MAY FUNCTION AS PART OF A SIGNALING PATHWAY THAT ACTS
CC       TO REGULATE THE SIZE OF THE BODY FAT DEPOT. AN INCREASE IN THE
CC       LEVEL OF OB MAY ACT DIRECTLY OR INDIRECTLY ON THE CNS TO INHIBIT
CC       FOOD INTAKE AND/OR REGULATE ENERGY EXPENDITURE AS PART OF A
CC       HOMEOSTATIC MECHANISM TO MAINTAIN CONSTANCY OF THE ADIPOSE MASS.
CC   -!- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE LEPTIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U58492; AAC50730.1; -.
DR   InterPro; IPR000065; -.
DR   Pfam; PF02024; Leptin; 1.
DR   PRINTS; PR00495; LEPTIN.
KW   Obesity; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    167       LEPTIN.
FT   DISULFID   117    167       BY SIMILARITY.
SQ   SEQUENCE   167 AA;  18953 MW;  E7D9F30628A5BBE9 CRC64;


  Query Match             72.3%;  Score 34;  DB 1;  Length 167;
  Best Local Similarity   85.7%;  Pred. No. 30;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          ||||| |
Db    116 SCHLPLA 122


RESULT  10
CGD3_ARATH
ID   CGD3_ARATH     STANDARD;      PRT;   376 AA.
AC   P42753; O49489;
DT   01-NOV-1995 (Rel. 32, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYCLIN DELTA-3.
GN   CYCD3 OR AT4G34160 OR F28A23.80.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. LANDSBERG ERECTA; TISSUE=Seedling;
RX   MEDLINE=95210930; PubMed=7696881;
RA   Soni R., Carmichael J.P., Shah Z.H., Murray J.A.H.;
RT   "A family of cyclin D homologs from plants differentially controlled
RT   by growth regulators and containing the conserved retinoblastoma
RT   protein interaction motif.";
RL   Plant Cell 7:85-103(1995).
RN   [2]
```

```
RP   REVISION TO 371.
RA   Murray J.A.H.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=20083488; PubMed=10617198;
RA   Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA   Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA   Harris B., Ansorge W., Brandt P., Grivell L., Rieger M.,
RA   Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA   Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T.,
RA   Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA   Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P.,
RA   Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA   Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA   Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA   Weitzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
RA   Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA   Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA   Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA   De Keyser A., Buysshaert C., Gielen J., Villarroel R., De Clercq R.,
RA   Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA   Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
RA   Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA   Borkova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA   Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA   Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA   Massenet O., Quigley F., Clabauld G., Muendlein A., Felber R.,
RA   Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA   Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA   Gibbons T., Weber N., Vandenbol M., Bargues M., Terol J., Torres A.,
RA   Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA   Heijnen L., Schwarz S., Scholler P., Heber S., Francs P., Bielke C.,
RA   Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA   Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA   Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA   Sehkon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA   Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA   Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA   Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA   Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA   Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA   Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C.,
RA   Antonoiu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA   Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA   Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA   Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA   Chen E., Marra M., Martienssen R., McCombie W.R.;
RT   "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 402:769-777(1999).
CC   -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X83371; CAA58287.1; -.
DR   EMBL; AL021961; CAA17556.1; -.
DR   EMBL; AL161584; CAB80133.1; -.
DR   InterPro; IPR000553; -.
DR   Pfam; PF00134; cyclin; 1.
DR   PROSITE; PS00292; CYCLINS; 1.
KW   Cyclin; Cell cycle; Cell division; Multigene family.
FT   CONFLICT    288    288       C -> G (IN REF. 3).
SQ   SEQUENCE   376 AA;  42747 MW;  FB8D5B6BC435FAC2 CRC64;
```

```
  Query Match             72.3%;  Score 34;  DB 1;  Length 376;
  Best Local Similarity   100.0%;  Pred. No. 59;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 HLPWA 7
          |||||
Db     367 HLPWA 371


RESULT  11
ACRO_HUMAN
ID   ACRO_HUMAN     STANDARD;      PRT;   421 AA.
AC   P10323;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ACROSIN PRECURSOR (EC 3.4.21.10).
GN   ACR OR ACRS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Testis;
RX   MEDLINE=89153568; PubMed=2493394;
RA   Baba T., Watanabe K., Kashiwabara S.-I., Arai Y.;
RT   "Primary structure of human proacrosin deduced from its cDNA
RT   sequence.";
RL   FEBS Lett. 244:296-300(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leukocyte;
RX   MEDLINE=90306003; PubMed=2114285;
RA   Keime S., Adham I.M., Engel W.;
RT   "Nucleotide sequence and exon-intron organization of the human
RT   proacrosin gene.";
RL   Eur. J. Biochem. 190:195-200(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92331659; PubMed=1628652;
RA   Vazquez-Levin M.H., Reventos J., Gordon J.W.;
RT   "Molecular cloning, sequencing and restriction mapping of the genomic
RT   sequence encoding human proacrosin.";
RL   Eur. J. Biochem. 207:23-26(1992).
RN   [4]
RP   DISCUSSION ON ABOVE PAPER.
RA   Adham I.A., Spitzer U., Schloesser M., Kremling H., Keime S.,
RA   Engel W.;
RL   Eur. J. Biochem. 207:27-28(1992).
CC   -!- FUNCTION: ACROSIN IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.
CC       IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY, IT
CC       IS SYNTHESIZED IN A ZYMOGEN FORM, PROACROSIN AND STORED IN THE
CC       ACROSOME.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL
CC       CLEAVAGE ARG-XAA >> LYS-LYS >> LYS-XAA.
CC   -!- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Y00970; CAA68784.1; -.
DR   EMBL; X54017; CAA37964.1; -.
DR   EMBL; X54018; CAA37964.1; JOINED.
```

```
DR   EMBL; X54019; CAA37964.1; JOINED.
DR   EMBL; X54020; CAA37964.1; JOINED.
DR   EMBL; M77378; AAA51572.1; -.
DR   EMBL; M77379; AAA51573.1; -.
DR   EMBL; M77380; AAA51574.1; -.
DR   EMBL; M77381; AAA51575.1; -.
DR   EMBL; X66188; CAA46956.1; -.
DR   EMBL; X54018; CAA46956.1; JOINED.
DR   EMBL; X54019; CAA46956.1; JOINED.
DR   EMBL; X54020; CAA46956.1; JOINED.
DR   PIR; S03330; S03330.
DR   PIR; S11674; S11674.
DR   PIR; S12063; S12063.
DR   PIR; S23499; S23499.
DR   MEROPS; S01.223; -.
DR   MIM; 102480; -.
DR   InterPro; IPR001254; -.
DR   InterPro; IPR001314; -.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.
FT   SIGNAL       1     19
FT   CHAIN       20    421       ACROSIN.
FT   CHAIN       20     42       ACROSIN LIGHT CHAIN.
FT   CHAIN       43      ?       ACROSIN HEAVY CHAIN.
FT   PROPEP       ?    421       PRO-RICH.
FT   DISULFID    25    154       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    29    162       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    73     89       BY SIMILARITY.
FT   DISULFID   177    246       BY SIMILARITY.
FT   DISULFID   209    225       BY SIMILARITY.
FT   DISULFID   236    266       BY SIMILARITY.
FT   CARBOHYD    22     22       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   210    210       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   ACT_SITE    88     88       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   142    142       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   240    240       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   CONFLICT    64     64       R -> T (IN REF. 1).
FT   CONFLICT   120    120       L -> V (IN REF. 1).
FT   CONFLICT   166    166       F -> L (IN REF. 1).
FT   CONFLICT   268    268       L -> R (IN REF. 1).
FT   CONFLICT   345    345       P -> R (IN REF. 1).
SQ   SEQUENCE   421 AA;  45799 MW;  62EE47DC25B4FB5D CRC64;
```

```
  Query Match             72.3%;  Score 34;  DB 1;  Length 421;
  Best Local Similarity   83.3%;  Pred. No. 65;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPW 6
           | ||||
Db     321 SAHLPW 326
```

```
RESULT  12
PABB_ECOLI
ID   PABB_ECOLI     STANDARD;      PRT;   453 AA.
AC   P05041;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PARA-AMINOBENZOATE SYNTHASE COMPONENT I (EC 4.1.3.-) (ADC SYNTHASE).
GN   PABB.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84239604; PubMed=6330050;
RA   Goncharoff P., Nichols B.P.;
RT   "Nucleotide sequence of Escherichia coli pabB indicates a common
RT   evolutionary origin of p-aminobenzoate synthetase and anthranilate
RT   synthetase.";
RL   J. Bacteriol. 159:57-62(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251358; PubMed=9097040;
RA   Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T.,
RA   Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T.,
RA   Saito N., Sampei G., Seki Y., Sivasundaram S., Tagami H.,
RA   Takeda J., Takemoto K., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 40.1-50.0 min region on the linkage map.";
RL   DNA Res. 3:379-392(1996).
RN   [4]
RP   SEQUENCE OF 42-377 FROM N.A.
RC   STRAIN=ECOR8, ECOR16, AND ECOR10;
RX   MEDLINE=95203706; PubMed=7896119;
RA   Guttman D.S., Dykhuizen D.E.;
RT   "Detecting selective sweeps in naturally occurring Escherichia coli.";
RL   Genetics 138:993-1003(1994).
CC   -!- FUNCTION: CATALYZES THE BIOSYNTHESIS OF 4-AMINO-4-DEOXYCHORISMATE
CC       (ADC) FROM CHORISMATE AND GLUTAMINE.
CC   -!- PATHWAY: FOLATE BIOSYNTHESIS PATHWAY. FIRST STEP IN THE
CC       BIOSYNTHESIS OF P-AMINOBENZOATE (PABA).
CC   -!- SUBUNIT: CONSISTS OF TWO NONIDENTICAL CHAINS: COMPONENT I
CC       CATALYZES THE FORMATION OF ADC BY BINDING CHORISMATE AND AMMONIA;
CC       COMPONENT II PROVIDES THE GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02673; AAA24266.1; -.
DR   EMBL; AE000275; AAC74882.1; -.
DR   EMBL; D90825; BAA15619.1; -.
DR   EMBL; U07762; AAC43282.1; -.
DR   EMBL; U07748; AAC43269.1; -.
DR   EMBL; U07749; AAC43270.1; -.
DR   PIR; A30251; AGEC1.
DR   EcoGene; EG10683; pabB.
DR   InterPro; IPR000350; -.
DR   Pfam; PF00425; chorismate_bind; 1.
DR   PRINTS; PR00095; ANTSNTHASEI.
KW   Lyase; Folate biosynthesis.
SQ   SEQUENCE   453 AA;  50969 MW;  DBF17DD5E17289D8 CRC64;
```

```
  Query Match             72.3%;  Score 34;  DB 1;  Length 453;
  Best Local Similarity   100.0%;  Pred. No. 69;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 HLPWA 7
           |||||
```

```
Db      27 HLPWA 31


RESULT  13
PABB_SALTY
ID   PABB_SALTY      STANDARD;      PRT;   454 AA.
AC   P12680;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PARA-AMINOBENZOATE SYNTHASE COMPONENT I (EC 4.1.3.-) (ADC SYNTHASE).
GN   PABB.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89056707; PubMed=3057324;
RA   Goncharoff P., Nichols B.P.;
RT   "Evolution of aminobenzoate synthases: nucleotide sequences of
RT   Salmonella typhimurium and Klebsiella aerogenes pabB.";
RL   Mol. Biol. Evol. 5:531-548(1988).
CC   -!- FUNCTION: CATALYZES THE BIOSYNTHESIS OF 4-AMINO-4-DEOXYCHORISMATE
CC       (ADC) FROM CHORISMATE AND GLUTAMINE.
CC   -!- PATHWAY: FOLATE BIOSYNTHESIS PATHWAY. FIRST STEP IN THE
CC       BIOSYNTHESIS OF P-AMINOBENZOATE (PABA).
CC   -!- SUBUNIT: CONSISTS OF TWO NONIDENTICAL CHAINS: COMPONENT I
CC       CATALYZES THE FORMATION OF ADC BY BINDING CHORISMATE AND AMMONIA;
CC       COMPONENT II PROVIDES THE GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M22079; AAA88618.1; -.
DR   PIR; A31132; A31132.
DR   StyGene; SG10274; pabB.
DR   InterPro; IPR000350; -.
DR   Pfam; PF00425; chorismate_bind; 1.
DR   PRINTS; PR00095; ANTSNTHASEI.
KW   Lyase; Folate biosynthesis.
SQ   SEQUENCE   454 AA;  50978 MW;  430B3949B4904546 CRC64;


  Query Match             72.3%;  Score 34;  DB 1;  Length 454;
  Best Local Similarity   100.0%;  Pred. No. 69;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 HLPWA 7
          |||||
Db     28 HLPWA 32


RESULT  14
WRN_HUMAN
ID   WRN_HUMAN       STANDARD;       PRT;  1432 AA.
AC   Q14191;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   WERNER SYNDROME HELICASE.
GN   WRN OR RECQL3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96181115; PubMed=8602509;
RA   Yu C.-E., Oshima J., Fu Y.-H., Wijsman E.M., Hisama F., Alisch R.,
RA   Matthews S., Nakura J., Miki T., Ouais S., Martin G.M., Mulligan J.,
RA   Schellenberg G.D.;
RT   "Positional cloning of the Werner's syndrome gene.";
RL   Science 272:258-262(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Paeper B.W., Gayle M., Brady W., Swartz A., Gillett L.A., Alisch R.S.,
RA   Mulligan J., Galas D., Fu Y.-H.;
RT   "Genomic structure of the human Werner's gene and cloning of the
RT   mouse homolog.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SUBCELLULAR LOCATION.
RX   MEDLINE=98284027; PubMed=9618508;
RA   Marciniak R.A., Lombard D.B., Johnson F.B., Guarente L.;
RT   "Nucleolar localization of the Werner syndrome protein in human
RT   cells.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:6887-6892(1998).
RN   [4]
RP   REVIEW ON VARIANTS.
RX   MEDLINE=99235545; PubMed=10220139;
RA   Moser M.J., Oshima J., Monnat R.J. Jr.;
RT   "WRN mutations in Werner syndrome.";
RL   Hum. Mutat. 13:271-279(1999).
RN   [5]
RP   VARIANT ARG-1367.
RX   MEDLINE=97173161; PubMed=9021029;
RA   Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
RA   Ikegami H., Higaki J., Edland S.D., Martin G.M., Ogihara T.;
RT   "Association of a polymorphic variant of the Werner helicase gene with
RT   myocardial infarction in a Japanese population.";
RL   Am. J. Med. Genet. 68:494-498(1997).
RN   [6]
RP   ERRATUM.
RA   Ye L., Miki T., Nakura J., Oshima J., Kamino K., Rakugi H.,
RA   Ikegami H., Higaki J., Edland S.D., Martin G.M., Ogihara T.;
RL   Am. J. Med. Genet. 70:103-103(1997).
RN   [7]
RP   VARIANTS ILE-387 AND LEU-1074.
RX   MEDLINE=98111850; PubMed=9450180;
RA   Meisslitzer C., Ruppitsch W., Weirich-Schwaiger H., Weirich H.G.,
RA   Jabkowsky J., Klein G., Schweiger M., Hirsch-Kauffmann M.;
RT   "Werner syndrome: characterization of mutations in the WRN gene in an
RT   affected family.";
RL   Eur. J. Hum. Genet. 5:364-370(1997).
RN   [8]
RP   VARIANT ILE-387.
RA   Vidal V., Bay J.-O., Champomier F., Grancho M., Beauville L.,
RA   Glowaczower C., Lemery D., Ferrara M., Bignon Y.-J.;
RT   "The 1396del A mutation and a missense mutation or a rare polymorphism
RT   of the WRN gene detected in a French Werner family with a severe
RT   phenotype and a case of an unusual vulvar cancer.";
RL   Hum. Mutat. 11:413-414(1998).
RN   [9]
RP   VARIANTS ALA-324 AND ARG-1367.
RX   MEDLINE=99167244; PubMed=10069711;
RA   Castro E., Ogburn C.E., Hunt K.E., Tilvis R., Louhija J.,
RA   Penttinen R., Erkkola R., Panduro A., Riestra R., Piussan C.,
RA   Deeb S.S., Wang L., Edland S.D., Martin G.M., Oshima J.;
RT   "Polymorphisms at the Werner locus: I. Newly identified polymorphisms,
RT   ethnic variability of 1367Cy/Arg, and its stability in a population
RT   of Finnish centenarians.";
RL   Am. J. Med. Genet. 82:399-403(1999).
CC   -!- FUNCTION: MAY BE INVOLVED IN THE CONTROL OF GENOMIC STABILITY.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC   -!- DISEASE: DEFECTS IN WRN ARE THE CAUSE OF WERNER SYNDROME (WS); A
CC       RARE AUTOSOMAL RECESSIVE PROGEROID SYNDROME CHARACTERIZED BY THE
CC       PREMATURE ONSET OF MULTIPLE AGE-RELATED DISORDERS, INCLUDING
CC       ATHEROSCLEROSIS, CANCER, NON-INSULIN-DEPENDENT DIABETES MELLITUS
```

```
CC       (NIDDM), OCULAR CATARACTS AND OSTEOPOROSIS. THE MAJOR CAUSE OF
CC       DEATH (AT A MEDIAN AGE OF 47) IS MYOCARDIAL INFARCTION (MI).
CC   -!- SIMILARITY: BELONGS TO THE RECQ SUBFAMILY OF HELICASES.
CC   -!- DATABASE: NAME=WRN; NOTE=WRN mutation db (Warner disease);
CC       WWW="http://www.pathology.washington.edu/werner/ws_wrn.html".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L76937; AAC41981.1; -.
DR   EMBL; AF091214; AAC63361.1; -.
DR   MIM; 604611; -.
DR   MIM; 277700; -.
DR   InterPro; IPR001410; -.
DR   InterPro; IPR001650; -.
DR   InterPro; IPR002121; -.
DR   InterPro; IPR002562; -.
DR   Pfam; PF01612; 3_5_exonuclease; 1.
DR   Pfam; PF00270; DEAD; 1.
DR   Pfam; PF00570; HRDC; 1.
DR   Pfam; PF00271; helicase_C; 1.
KW   Helicase; ATP-binding; Nuclear protein; Polymorphism.
FT   DOMAIN      507    510       POLY-GLU.
FT   NP_BIND     571    578       ATP (BY SIMILARITY).
FT   SITE        668    671       DEAH BOX.
FT   DOMAIN     1150   1229       HRDC.
FT   VARIANT     324    324       T -> A.
FT                                /FTId=VAR_006904.
FT   VARIANT     387    387       M -> I.
FT                                /FTId=VAR_006905.
FT   VARIANT    1074   1074       F -> L.
FT                                /FTId=VAR_007903.
FT   VARIANT    1367   1367       C -> R (ASSOCIATED WITH A HIGHER RISK OF
FT                                MYOCARDIAL INFARCTION).
FT                                /FTId=VAR_006906.
SQ   SEQUENCE   1432 AA;  162494 MW;  9F02C0059F7B62EB CRC64;
```

```
  Query Match             72.3%;  Score 34;  DB 1;  Length 1432;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 SCHLPWA 7
          |||: ||
Db    863 SCHVLWA 869
```

```
RESULT  15
YJJW_ECOLI
ID   YJJW_ECOLI      STANDARD;      PRT;   287 AA.
AC   P39409;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL 31.5 KDA PROTEIN IN OSMY-DEOC INTERGENIC REGION.
GN   YJJW.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=95334362; PubMed=7610040;
RA   Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA   Blattner F.R.;
RT   "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT   region from 92.8 through 100 minutes.";
RL   Nucleic Acids Res. 23:2105-2119(1995).
CC   -!- COFACTOR: BINDS A 4FE-4S CLUSTER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ORGANIC RADICAL ACTIVATING ENZYMES
CC       FAMILY.
CC   -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC       'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U14003; AAA97275.1; -.
DR   EMBL; AE000508; AAC77332.1; -.
DR   HSSP; P00198; 1FDN.
DR   EcoGene; EG12599; yjjW.
DR   InterPro; IPR001450; -.
DR   InterPro; IPR001989; -.
DR   Pfam; PF00037; fer4; 1.
DR   PROSITE; PS00198; 4FE4S_FERREDOXIN; 2.
DR   PROSITE; PS01087; RADICAL_ACTIVATING; 1.
KW   Hypothetical protein; Iron-sulfur; 4Fe-4S.
FT   METAL       31     31       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       35     35       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       38     38       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       47     47       IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT   METAL       50     50       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL       53     53       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL       57     57       IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT   METAL       76     76       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       79     79       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       82     82       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT   METAL       86     86       IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
SQ   SEQUENCE   287 AA;  31490 MW;  E08BB429519E54B3 CRC64;
```

```
  Query Match             70.2%;  Score 33;  DB 1;  Length 287;
  Best Local Similarity   66.7%;  Pred. No. 69;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 SCHLPW 6
          :|| ||
Db     37 NCHNPW 42
```

Search completed: July 3, 2001, 20:51:07
Job time: 206 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: July 3, 2001, 20:46:35 ; Search time 18.83 Seconds
(without alignments)
28.318 Million cell updates/sec

Title: US-09-377-081-18
Perfect score: 47
Sequence: 1 SCHLPWA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	47	100.0	166	2	I53166	leptin precursor -
2	47	100.0	167	1	LTHU	leptin precursor -
3	39	83.0	773	2	T00502	probable receptor-
4	37	78.7	1401	2	T17452	Werner syndrome pr
5	37	78.7	1401	2	T30247	Werner syndrome pr
6	36	76.6	367	2	T01751	gibberellin 20-oxi
7	36	76.6	397	2	H84578	probable RING zinc
8	36	76.6	6805	2	S20901	titin - rabbit (fr
9	36	76.6	26926	1	I38344	titin, cardiac mus
10	35	74.5	409	1	JS0759	membrane dipeptida
11	35	74.5	411	1	S29848	membrane dipeptida
12	34	72.3	180	2	T34745	probable proteinas
13	34	72.3	258	2	T27393	hypothetical prote
14	34	72.3	291	2	T36190	probable transfera
15	34	72.3	293	2	T31146	probable DNA inver
16	34	72.3	317	2	G83544	probable transcrip
17	34	72.3	376	2	T05420	cyclin delta-3 - A
18	34	72.3	384	2	G83040	hypothetical prote
19	34	72.3	421	1	S11674	acrosin (EC 3.4.21
20	34	72.3	425	2	D83186	hypothetical prote
21	34	72.3	453	1	AGEC1	p-aminobenzoate sy
22	34	72.3	453	2	E85792	p-aminobenzoate sy
23	34	72.3	454	2	A31132	p-aminobenzoate sy
24	34	72.3	538	2	S76175	hypothetical prote
25	34	72.3	540	2	B47417	insulin receptor-r
26	34	72.3	1155	2	G96539	hypothetical prote
27	33	70.2	165	2	C48232	cysteine-rich exte
28	33	70.2	211	2	S12252	self incompatibili
29	33	70.2	221	2	T07176	extensin homolog -
30	33	70.2	287	1	S56603	probable pyruvate
31	33	70.2	287	2	D86137	probable activatin
32	33	70.2	290	2	T16540	hypothetical prote
33	33	70.2	304	2	G83435	probable DNA methy
34	33	70.2	320	2	T35265	probable D-amino a
35	33	70.2	388	2	H65126	probable general s
36	33	70.2	504	2	G82631	glutamine syntheta
37	33	70.2	531	2	A35343	glucuronosyltransf
38	33	70.2	570	2	F70844	probable fusion pr
39	33	70.2	579	4	D40201	artifact-warning s
40	33	70.2	622	2	D84493	probable retroelem
41	33	70.2	628	2	T02602	vacuolar sorting r
42	33	70.2	628	2	T02604	probable vacuolar
43	33	70.2	665	2	T18979	hypothetical prote
44	33	70.2	666	2	E71565	probable glycogen
45	33	70.2	666	2	G81717	glycosyl hydrolase

ALIGNMENTS

RESULT 1
I53166
leptin precursor - human
N;Alternate names: obese
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 16-Jul-1999
C;Accession: I53166; G02328
R;Masuzaki, H.; Ogawa, Y.; Isse, N.; Satoh, N.; Okazaki, T.; Shigemoto, M.; Mori, K.
Diabetes 44, 855-858, 1995
A;Title: Human obese gene expression. Adipocyte-specific expression and regional dif
A;Reference number: I53166; MUID:95309556
A;Accession: I53166
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-166 <RES>
A;Cross-references: GB:D49487; NID:g904211; PIDN:BAA08448.1; PID:g904212
R;Chehab, F.F.; Lim, M.E.
submitted to the EMBL Data Library, December 1995
A;Reference number: H01063
A;Accession: G02328
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-166 <CHE>
A;Cross-references: EMBL:U43415; NID:g1163105; PIDN:AAC31660.1; PID:g1163106
C;Genetics:
A;Gene: GDB:LEP; OB; OBS
A;Cross-references: GDB:136420; OMIM:164160
A;Map position: 7q32.1-7q32.1
A;Introns: 48/3
C;Superfamily: leptin

Query Match 100.0%; Score 47; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 115 SCHLPWA 121

RESULT 2
LTHU
leptin precursor - human
N;Alternate names: obese protein; obesity factor
C;Species: Homo sapiens (man)
C;Date: 28-Jul-1995 #sequence_revision 16-Aug-1996 #text_change 01-Dec-2000
C;Accession: A38952; JE0148
R;Zhang, Y.; Proenca, R.; Maffei, M.; Barone, M.; Leopold, L.; Friedman, J.M.
Nature 372, 425-432, 1994
A;Title: Positional cloning of the mouse obese gene and its human homologue.
A;Reference number: S50863; MUID:95075453

A;Accession: A38952
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-167 <ZHA>
A;Cross-references: GB:U18915; NID:g623331; PIDN:AAA60470.1; PID:g623332
R;Liao, H.J.; Deng, Y.B.; Chen, X.M.; Ye, Y.Z.
Chinese Biochem. J. 13, 249-253, 1997
A;Title: Cloning of chinese obesity gene and construction of prokaryotic expression vect
A;Reference number: JE0148
A;Accession: JE0148
A;Molecule type: mRNA
A;Residues: 'M',22-167 <LIA>
A;Experimental source: adipose
A;Note: the author translated GAC for residue 148 as Ser
C;Genetics:
A;Gene: GDB:LEP; OB; OBS
A;Cross-references: GDB:136420; OMIM:164160
A;Map position: 7q31.3-7q31.3
C;Superfamily: leptin
C;Keywords: adipose tissue
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-167/Product: leptin #status predicted <MAT>

Query Match 100.0%; Score 47; DB 1; Length 167;
Best Local Similarity 100.0%; Pred. No. 0.43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||||||
Db 116 SCHLPWA 122

RESULT 3
T00502
probable receptor-like protein kinase At2g23300 [imported] - Arabidopsis thaliana
N;Alternate names: protein kinase homolog T20D16.7
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 23-Mar-2001
C;Accession: T00502; A84623
R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, November 1997
A;Description: Arabidopsis thaliana chromosome II BAC T20D16 genomic sequence.
A;Reference number: Z14164
A;Accession: T00502
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-773 <ROU>
A;Cross-references: EMBL:AC002391; NID:g2642427; PID:g2642433
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: A84623
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-773 <STO>
A;Cross-references: GB:AE002093; NID:g2642433; PIDN:AAB87101.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g23300; T20D16.7
A;Map position: 2
A;Introns: 545/1
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match 83.0%; Score 39; DB 2; Length 773;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHLPW 6
|||||
Db 552 CHLPW 556

RESULT 4
T17452
Werner syndrome protein - mouse
N;Alternate names: Wrn protein
C;Species: Mus musculus (house mouse)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17452
R;Paeper, B.W.; Gayle, M.; Brady, W.; Swartz, A.; Gillett, L.A.; Alisch, R.S.; Mullig
submitted to the EMBL Data Library, September 1998
A;Description: Genomic structure of the human Werner's gene and cloning of its mouse
A;Reference number: Z18794
A;Accession: T17452
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1401 <PAE>
A;Cross-references: EMBL:AF091215; NID:g3885837; PID:g3885838; PIDN:AAC78077.1
C;Genetics:
A;Gene: Wrn

Query Match 78.7%; Score 37; DB 2; Length 1401;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||| ||
Db 828 SCHLLWA 834

RESULT 5
T30247
Werner syndrome protein typeI - mouse
N;Alternate names: Wrn typeI protein
C;Species: Mus musculus (house mouse)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C;Accession: T30247
R;Imamura, O.; Ichikawa, K.; Yamabe, Y.; Goto, M.; Sugawara, M.; Furuichi, Y.
Genomics 41, 298-300, 1997
A;Title: Cloning of a mouse homologue of the human Werner syndrome gene and assignmen
A;Reference number: Z20785; MUID:97288537
A;Accession: T30247
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1401 <IMA>
A;Cross-references: EMBL:D86526; NID:g2130972; PIDN:BAA20269.1; PID:g2130973
A;Experimental source: strain BALB/c; testis/spleen
C;Genetics:
A;Gene: WRN typeI
A;Map position: 8A4

Query Match 78.7%; Score 37; DB 2; Length 1401;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||| ||
Db 828 SCHLLWA 834

RESULT 6
T01751
gibberellin 20-oxidase - common tobacco
N;Alternate names: Ntc16 protein
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 20-Jun-2000
C;Accession: T01751
R;Tanaka-Ueguchi, M.; Itoh, H.; Oyama, N.; Koshioka, M.; Matsuoka, M.
submitted to the EMBL Data Library, July 1998

A;Description: Over-expression of a tobacco homeobox gene, NTH15, decreases the expressi
A;Reference number: Z14418
A;Accession: T01751
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-367 <TAN>
A;Cross-references: EMBL:AB016084
C;Genetics:
A;Gene: Ntcl6
C;Superfamily: 1-aminocyclopropane-1-carboxylate oxidase

Query Match 76.6%; Score 36; DB 2; Length 367;
Best Local Similarity 83.3%; Pred. No. 58;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SCHLPW 6
||:|||
Db 133 SCNLPW 138

RESULT 7
H84578
probable RING zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: H84578
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487
A;Accession: H84578
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-397 <STO>
A;Cross-references: GB:AE002093; NID:g4191790; PIDN:AAD10159.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g19610
A;Map position: 2

Query Match 76.6%; Score 36; DB 2; Length 397;
Best Local Similarity 80.0%; Pred. No. 62;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CHLPW 6
||:||
Db 332 CHVPW 336

RESULT 8
S20901
titin - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C;Accession: S20901; I46520
R;Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A;Title: Towards a molecular understanding of titin.
A;Reference number: S20897; MUID:92258380
A;Accession: S20901
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-6805 <LAB>
A;Cross-references: EMBL:X64696
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
R;Labeit, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Hsieh, C.L.; Francke, U.;
Nature 345, 273-276, 1990
A;Title: A regular pattern of two types of 100-residue motif in the sequence of titin.
A;Reference number: I46520; MUID:90238553
A;Accession: I46520
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 4235-5250 <LA2>
A;Cross-references: EMBL:X17329; NID:g1756; PIDN:CAA35207.1; PID:g930251
C;Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology;
C;Keywords: muscle

Query Match 76.6%; Score 36; DB 2; Length 6805;
Best Local Similarity 71.4%; Pred. No. 6.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||: ||
Db 576 SCHVSWA 582

RESULT 9
I38344
titin, cardiac muscle [validated] - human
N;Alternate names: connectin
N;Contains: serine/threonine-specific protein kinase (EC 2.7.1.-)
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Sep-2000
C;Accession: I38344; I38345; S20898; S20897; S20899; S63665; S37393
R;Labeit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A;Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A;Reference number: A57430; MUID:96026330
A;Accession: I38344
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB,
A;Molecule type: mRNA
A;Residues: 1-26926 <LAB1>
A;Cross-references: EMBL:X90568; NID:g1017424; PID:g1017425
R;Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.
Biochemistry 34, 553-561, 1995
A;Title: Dissecting titin into its structural motifs: identification of an alpha-hel:
A;Reference number: I38345; MUID:95119041
A;Accession: I38345
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1977-2014 <MUS>
A;Cross-references: EMBL:X83270; NID:g602579; PIDN:CAA58243.1; PID:g602580
A;Note: conformation and properties are reported for a synthetic peptide correspondir
R;Labeit, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A;Title: Towards a molecular understanding of titin.
A;Reference number: S20897; MUID:92258380
A;Accession: S20898
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 13597-14200,'I',14202-14696 <LAB2>
A;Cross-references: EMBL:X64698; NID:g37192; PIDN:CAA45939.1; PID:g37193
A;Accession: S20897
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 16330-16382,'S',16384-16756,'F',16758-16860 <LAB3>
A;Cross-references: EMBL:X64699; NID:g37190; PIDN:CAA45940.1; PID:g37191
A;Accession: S20899
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 'P',22278-22431,'R',22433-22448,'G',22450-22453,'Q',22455-22480,'TR',2248
A;Cross-references: EMBL:X64697; NID:g37190; PIDN:CAA45938.1; PID:g37195
R;Kolmerer, B.; Olivieri, N.; Witt, C.C.; Herrmann, B.G.; Labeit, S.
J. Mol. Biol. 256, 556-563, 1996
A;Title: Genomic organization of M line titin and its tissue-specific expression in t
A;Reference number: S63665; MUID:96177761
A;Accession: S63665
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 26729-26825 <KOL>
A;Cross-references: EMBL:X92412; NID:g1236761
R;Gautel, M.; Leonard, K.; Labeit, S.

EMBO J. 12, 3827-3834, 1993
A;Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiat
A;Reference number: S37393; MUID:94008990
A;Accession: S37393
A;Molecule type: mRNA
A;Residues: 26831-26926 <GAU>
R;Improta, S.; Politou, A.S.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, February 1996
A;Reference number: A66736; PDB:1TIT
A;Contents: annotation; conformation by (1)H-NMR, residues 5253-5341
R;Pfuhl, M.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, August 1996
A;Reference number: A66201; PDB:1NCT
A;Contents: annotation; conformation by (1)H-NMR, residues 'S',26059-26155
C;Genetics:
A;Gene: GDB:TTN
A;Cross-references: GDB:127867; OMIM:188840
A;Map position: 2q31-2q32
C;Function:
A;Description: structural protein forming filaments in striated muscle
C;Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; prc
C;Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glycc
structural protein
F;24752-25008/Domain: protein kinase homology <KIN>
F;84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,
98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1
tatus predicted
F;16780,16976,17579;17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18
,21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248
F;26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 76.6%; Score 36; DB 1; Length 26926;
Best Local Similarity 71.4%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||: ||
Db 12217 SCHVSWA 12223

RESULT 10
JS0759
membrane dipeptidase (EC 3.4.13.19) precursor - pig
N;Alternate names: renal dipeptidase
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jan-2000
C;Accession: JS0759; PS0394; S13059; S08194
R;Satoh, S.; Koyama, S.; Ohtuka, K.; Keida, Y.; Niwa, M.; Kohsaka, M.
submitted to JIPID, September 1992
A;Reference number: JS0759
A;Accession: JS0759
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-409 <SAT>
A;Cross-references: DDBJ:D13142; NID:g217704; PIDN:BAA02433.1; PID:g217705
A;Experimental source: kidney
A;Accession: PS0394
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 75-129,'Q',131-312,'L',314-409 <SA2>
A;Cross-references: DDBJ:D13143; NID:g217703
A;Experimental source: kidney
R;Rached, E.; Hooper, N.M.; James, P.; Semenza, G.; Turner, A.J.; Mantei, N.
Biochem. J. 271, 755-760, 1990
A;Title: cDNA cloning and expression in Xenopus laevis oocytes of pig renal dipeptidase,
A;Reference number: S13059; MUID:91058511
A;Accession: S13059
A;Molecule type: mRNA
A;Residues: 1-409 <RAC>
A;Cross-references: EMBL:X53730; NID:g2101; PIDN:CAA37762.1; PID:g2102
A;Note: parts of this sequence, including the amino end of the mature protein, were dete
A;Note: the authors refer to the old number EC 3.4.13.11

R;Hooper, N.M.; Keen, J.N.; Turner, A.J.
Biochem. J. 265, 429-433, 1990
A;Title: Characterization of the glycosyl-phosphatidylinositol-anchored human renal d
A;Reference number: S08193; MUID:90147607
A;Accession: S08194
A;Molecule type: protein
A;Residues: 17-39 <HOO>
C;Complex: homodimer, disulfide linked
C;Function:
A;Description: hydrolyzes a broad range of dipeptides; hydolyzes leukotriene D4 to le
A;Note: has been implicated in the renal metabolism of glutathione conjugates
C;Superfamily: membrane dipeptidase
C;Keywords: blocked carboxyl end; dipeptide hydrolase; glycoprotein; homodimer; kidne
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-384/Product: membrane dipeptidase #status predicted <MAT>
F;385-409/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;57/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;235,286,289/Binding site: zinc (His) #status predicted
F;279/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;384/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature f

Query Match 74.5%; Score 35; DB 1; Length 409;
Best Local Similarity 71.4%; Pred. No. 93;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SCHLPWA 7
|||: |||
Db 169 SCNTPWA 175

RESULT 11
S29848
membrane dipeptidase (EC 3.4.13.19) precursor - human
N;Alternate names: dehydropeptidase-I; renal dipeptidase
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jan-2000
C;Accession: S29848; JS0756; A35467; JS0757; S08193; PX0021
R;Satoh, S.; Kusunoki, C.; Konta, Y.; Niwa, M.; Kohsaka, M.
Biochim. Biophys. Acta 1172, 181-183, 1993
A;Title: Cloning and structural analysis of genomic DNA for human renal dipeptidase.
A;Reference number: S29848; MUID:93176806
A;Accession: S29848
A;Molecule type: DNA
A;Residues: 1-411 <SAT>
A;Cross-references: DDBJ:D13136; NID:g219597; DDBJ:D13137; NID:g219598; PIDN:BAA02430
A;Note: the authors translated the codon TCC for residue 9 as Pro
R;Satoh, S.; Kusunoki, C.; Konta, Y.; Niwa, M.; Kohsaka, M.
submitted to JIPID, September 1992
A;Reference number: JS0756
A;Accession: JS0756
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-411 <SA2>
A;Cross-references: DDBJ:D13128; NID:g219589
R;Adachi, H.; Tawaragi, Y.; Inuzuka, C.; Kubota, I.; Tsujimoto, M.; Nishihara, T.; Na
J. Biol. Chem. 265, 3992-3995, 1990
A;Title: Primary structure of human microsomal dipeptidase deduced from molecular clo
A;Reference number: A35467; MUID:90154088
A;Accession: A35467
A;Molecule type: mRNA
A;Residues: 1-8,'P',10-101,'R',103-124,'R',126-411 <ADA>
A;Cross-references: GB:J05257; NID:g598188; PIDN:AAB59410.1; PID:g598189
R;Satoh, S.; Ohtuka, K.; Kusunoki, C.; Konta, Y.; Keida, Y.; Niwa, M.; Kohsaka, M.
submitted to JIPID, September 1992
A;Reference number: JS0757
A;Accession: JS0757
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8,'P',10-144,'P',146-174,'S',176-411 <SA3>
A;Cross-references: DDBJ:D13138; NID:g219584; PIDN:BAA02431.1; PID:g219585
R;Hooper, N.M.; Keen, J.N.; Turner, A.J.
Biochem. J. 265, 429-433, 1990

```
A;Title: Characterization of the glycosyl-phosphatidylinositol-anchored human renal dipe
A;Reference number: S08193; MUID:90147607
A;Accession: S08193
A;Molecule type: protein
A;Residues: 17-35,'X',37-39 <HOO>
A;Experimental source: kidney
R;Adachi, H.; Kubota, I.; Okamura, N.; Iwata, H.; Tsujimoto, M.; Nakazato, H.; Nishihara
J. Biochem. 105, 957-961, 1989
A;Title: Purification and characterization of human microsomal dipeptidase.
A;Reference number: PX0021; MUID:89359222
A;Accession: PX0021
A;Molecule type: protein
A;Residues: 17-35,'G' <AD2>
A;Experimental source: kidney
R;Adachi, H.; Katayama, T.; Nakazato, H.; Tsujimoto, M.
Biochim. Biophys. Acta 1163, 42-48, 1993
A;Title: Importance of Glu-125 in the catalytic activity of human renal dipeptidase.
A;Reference number: A58391; MUID:93237320
A;Contents: annotation; active site by labeling and mutagenesis
C;Genetics:
A;Gene: GDB:DPEP1
A;Cross-references: GDB:128059; OMIM:179780
A;Map position: 16q24-16q24
A;Introns: 35/2; 79/3; 124/1; 174/2; 197/3; 256/3; 285/1; 310/2; 355/3
C;Complex: homodimer, disulfide linked
C;Function:
A;Description: hydrolyzes a broad range of dipeptides; hydolyzes leukotriene D4 to leuko
A;Note: has been implicated in the renal metabolism of glutathione conjugates
C;Superfamily: membrane dipeptidase
C;Keywords: blocked carboxyl end; dipeptide hydrolase; glycoprotein; homodimer; kidney;
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-384/Product: membrane dipeptidase #status predicted <MAT>
F;385-411/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;57,279,332,358/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;141/Active site: Glu #status predicted
F;235,286,289/Binding site: zinc (His) #status predicted
F;384/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form

  Query Match             74.5%;  Score 35;  DB 1;  Length 411;
  Best Local Similarity   71.4%;  Pred. No. 94;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           ||: |||
Db     169 SCNTPWA 175


RESULT  12
T34745
probable proteinase pfpI - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jan-2000
C;Accession: T34745
R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z21555
A;Accession: T34745
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-180 <SEE>
A;Cross-references: EMBL:AL033505; PIDN:CAA22052.1; GSPDB:GN00070; SCOEDB:SC1E6.24c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: pfpI; SCOEDB:SC1E6.24c
C;Superfamily: Archaeoglobus intracellular proteinase I

  Query Match             72.3%;  Score 34;  DB 2;  Length 180;
  Best Local Similarity   80.0%;  Pred. No. 69;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 CHLPW 6
           || ||
Db     106 CHAPW 110


RESULT  13
T27393
hypothetical protein Y75B8A.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Mar-2000
C;Accession: T27393
R;Barlow, K.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z20361
A;Accession: T27393
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-258 <WIL>
A;Cross-references: EMBL:AL033514; NID:e1343251; PIDN:CAA22093.1; CESP:Y75B8A.2
A;Experimental source: clone Y75B8A
C;Genetics:
A;Gene: CESP:Y75B8A.2
A;Introns: 63/2; 133/1; 226/3
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

  Query Match             72.3%;  Score 34;  DB 2;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 93;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 HLPWA 7
           |||||
Db     153 HLPWA 157


RESULT  14
T36190
probable transferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36190
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21600
A;Accession: T36190
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-291 <SAU>
A;Cross-references: EMBL:AL035707; PIDN:CAB38886.1; GSPDB:GN00070; SCOEDB:SCE29.11c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE29.11c

  Query Match             72.3%;  Score 34;  DB 2;  Length 291;
  Best Local Similarity   71.4%;  Pred. No. 1e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 SCHLPWA 7
           :| ||||
Db     275 ACALPWA 281


RESULT  15
T31146
probable DNA invertase - Sphingomonas aromaticivorans plasmid pNL1
C;Species: Sphingomonas aromaticivorans
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000
C;Accession: T31146
R;Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.
submitted to the EMBL Data Library, July 1998
A;Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas arom
A;Reference number: Z20992
```

A;Accession: T31146
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-293 <ROM>
A;Cross-references: EMBL:AF079317; NID:g3378261; PID:g3378287; PIDN:AAD03870.1
C;Genetics:
A;Genome: plasmid pNL1
A;Note: orf181

Query Match 72.3%; Score 34; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 HLPWA 7
||||| |
Db 181 HLPWA 185

Search completed: July 3, 2001, 20:48:10
Job time: 95 sec